CALL No. {516 / R733P ACC. NO. 17337

AUTHOR

TITLE A priori pathometry

THE BOOK MUST BE CHECKED AT THE TIME OF ISSUE

ALLAMA IQBAL LIBRARY
UNIVERSITY OF KASHMIR

Acc. No. ________ *Call No.* ________

1. This book should be returned on or before the last date stamped.
2. Overdue charges will be levied under rules for each day if the book is kept beyond the date stamped above.
3 Books lost, defaced or injured in any way shall have to be replaced by the borrower.

Help to keep this book fresh and clean

A PRIORI PATHOMETRY

BY

SIR RONALD ROSS
K.C.B., K.C.M.G., F.R.S., N.L.

AND

HILDA P. HUDSON
O.B.E., Sc.D., M.A.

PARTS I, II, and III

LONDON:
HARRISON AND SONS, LTD., 44, ST. MARTIN'S LANE, W.C. 2.
Printers in Ordinary and Booksellers to His Majesty

Price: Ten Shillings and Sixpence

1931

Paid

17337
18-10-57

ALLAMA IQBAL LIBRARY
17337

MATHEMATICAL WORKS

By SIR RONALD ROSS, K.C.B., K.C.M.G., F.R.S., N.L.
THE ROSS INSTITUTE, PUTNEY HEATH, LONDON, S.W.15

---

1. The Algebra of Space, being a brief Description of a System of Geometrical Algebra placed on an Arithmetical Basis
   G. Philip & Son, 32, Fleet Street, London, 1901, 31 pages, quarto. Combines the systems of Hamilton & Grassmann
2. Explicit Functional Notation with Remarks on the Solution of Equations by Operative Division
   Liverpool Mathematical Society, 25.4.1903 (*partly oral*)
3. Verb-Functions, with Notes on the Solution of Equations by Operative Division
   Proc. Roy. Irish Acad., Dublin, 1905, April, Sec. A, Vol. XXV, 3, pp. 31–76
4. A Method of Solving Algebraic Equations (By Iteration)
   Nature, London, 1908, Oct. 29
5. *Ibid.*
   Nature, London, 1909, Feb. 4
6. The Analysis of Equations by a Study of their Partial Roots, 1914–15
   Unfinished, mostly given in 7
7. The Solution of Equations by Operative Division
   Science Progress (Murray), London, 1915, Pt. I, Oct., pp. 218–236; 1916, Pt. II, Jan., pp. 393–411; and 1916, Pt. III, April, pp. 573–600
8. Operative Algebra: Operative Involution
   Ibid., 1918, October

9. On the Iteration of Multinomials, 1916
   Unpublished

10. The Iteration of Multinomials, 1926
    Unpublished

11. An Application of the Theory of Probabilities to the Study of *a priori* Pathometry
    Proc. Roy. Soc., London, 1916, A, Vol. 92, *Pt. I*
    Proc. Roy. Soc., London, 1917, A, Vol. 93, *Pts. II and III* (with Miss H. P. Hudson)

12. Isosceles Trigonometry
    Science Progress (Murray), London, 1919, Vol. XIII, Jan., pp. 485–6

13. The Operative Roots of the Circle-Function
    Science Progress (Murray), London, 1921, July, pp. 116–132

14. Constructive Epidemiology
    Brit. Med. Journ., London, 1929, April 13

15. Solid Space-Algebra
    Harrison & Sons, Ltd., London, 1929

16. Incomplete Works :—

1.	Operative Logarithms	27.10.20
2.	Operative Geometry, Operative Products ..	10.11.20
3.	Operative Geometry, Rotations	16.11.20
4.	Intrinsic Operations	18.11.20
5.	Operative Geometry, Transformations ..	18.12.20
6.	Operative Geometry, The Straight Line ..	18.12.20
7.	Trigonometrical Operations	5.12.20
8.	Two Unfinished Drafts ..	12.12.20 & 17.5.21
9.	Trigonometrical Operations, Incorrect ..	22.1.21
10.	Note on Operative Trigonometry	17.4.21
11.	The Genesis of Operations	21.4.21
12.	Operative Dynamics	8.1.23
13.	Point-Spheres	14.1.23
14.	An Introduction to Explicit Operations—	
	(1st draft)	3.2.23
	(2nd draft)	3.2.23

16. Incomplete Works (*continued*) :—

15. Rotary Equation of Curves	21.2.23
16. Relations of Co-ordinates	2.4.23
17. Rectilinear Figures (11 cont.)	12.4.23
18. Trigonometrical Formulae in K-terms	28.5.23
19. General Notes	8.2.24
20. Explicit Operations (3rd draft)	1929

17. The Solution of Equations by Iteration
    Harrison & Sons, Ltd., London, 1930

18. Iteration by Explicit Operations
    Harrison & Sons, Ltd., London, 1930

19. Circles by Explicit Operations (*forthcoming*)

20. *A priori* Pathometry
    Harrison & Sons, Ltd., London, 1931

21. An Outline of Solid Space-Algebra (*forthcoming*)

Four note books on Physics and Mathematics, commenced 1882
Two note books on Explicit Operations, commenced 1885

Folio I.—Algebra of Space and Explicit Operations
,, II.—A. General—1885, 1912
B. Roots by Iteration, 3.9.08–13.1.09
,, III.—General, 1908–1915

Two octavo note books, workings—
(1) 1912
(2) 1912–15

PREFACE

This book is a republication of a paper in three parts, of which the first part was written by me and the second and third parts by me in conjunction with Miss Hilda P. Hudson. They were published by the Royal Society in its 'Proceedings,' A, 1916 and 1917. Since that time, I have not been able to continue these studies.

The whole of the subject was also dealt with by the Finite Calculus at the end of my book "Prevention of Malaria," 2nd Edition, London, Murray, 1911, but this Addendum was written in such a hurry and under the press of so much other work, that it is not very clear.

The earlier chapters of the same book, especially Sections 26, 27 and 28, also contained an elementary kind of pathometry for non-mathematical readers, which does not pretend to be exact, but of which copies can be obtained from my Personal Secretary at this Institute.

I should add here that the studies contained in our original paper published by the Royal Society have been followed up by Prof. W. O. McCormack and Lt Colonel A. J. McKendrick in another paper in the 'Proceedings of the Royal Society,' A, Vol. 115, May, 1927. They used Integral Equations; and their work was favourably reviewed by Miss Hilda P. Hudson in 'Science Progress,' Vol. XXIII, January, 1929.

Metaxanous or diheteroxenic parasitism is not dealt with in this book, but was briefly considered in my "Prevention of Malaria," 2nd Edition, p. 665 (10 and 11) The analytical problem was solved by A. J. Lotka in 'Nature,' 8th Feb. 1912, p. 497, and elsewhere.

The subject is complicated by the fact recently discovered by L. W. Hackett in Italy, that *A. maculipennis*, usually a man-biting insect, often prefers domestic animals, when there need be no relation between the amount of malaria and the numbers of the carriers.

RONALD ROSS.

The Ross Institute,
Putney Heath,
London, S.W.15.

20th August, 1930.

A PRIORI PATHOMETRY
PART I

By Sir Ronald Ross, K.C.B., K.C.M.G., F.R.S., N.L.

Read November 11, 1915

CONTENTS

Part I

Section I

Page
(i) Introduction 5

Section II

(i) Statement of the Problem 9

Section III

(i) The Differential Equations 9
(ii) The Variation-Elements n, m, i, e, N, M, I, E 10
(iii) The Reversion-Element r 11
(iv) The Happening-Element h 12
(v) Independent and Dependent Happenings 12

Section IV

(i) Independent Happenings: h or F constant 13
(ii) The Equivariant Case: h constant: $e = V$ 13
(iii) Integrations 14
(iv) Analysis of x and f 15
(v) Integrals of P, Z, A, F, required for certain questions 17

Section V

Page
(i) Independent Happenings: h constant: $v \neq$ V 20
(ii) Integrations 20
(iii) Analysis of x 21
(iv) Analysis of f and P. Integral of P 22

Section VI

(i) Independent Happenings: F constant 23

Section VII

(i) Dependent Happenings. Proportional Happening: $h = cx$ 24
(ii) Integrations 25
(iii) Analysis of x 26
(iv) Analysis of dx/dt, and other matters 26
(v) Analysis of f 28
(vi) Further analysis of f and the constants K, L, c 30
(vii) Analysis of P 33
(viii) Integrals of P, Z, F 34
References 37

PART II

By Sir Ronald Ross and Hilda P. Hudson, O.B.E., M.A., Sc.D.

Read—February 1, 1917

Preface to Parts II and III 39

Section VIII

(i) Hypothetical Epidemics. Deduction of Constants 40
(ii) Examples with Table 42
(iii) Detailed Calculation of x and f 47

Section IX

(i) Hypometric Happening: KL negative 48
(ii) The Happening ceases: $c = 0$ 50
(iii) c lies between zero and D 51
(iv) c lies between D and D + R 51
(v) Isometric Happening: $c =$ D + R 52
(vi) Remarks 52

Section X

(i) Parameter Analysis of l and P 52

PART III

By Sir RONALD ROSS and HILDA P. HUDSON

Read—February 1, 1917

SECTION XI

PAGE
(i) Variable Happening : c varies 55
(ii) Discontinuous Change 55
(iii) Continuous Variation : $c = \phi'(t)$ 56
(iv) Graphical Treatment of x 57
(v) Graphical Treatment of f 60
(vi) f Assigned 61
(vii) More General Equations 65

SECTION XII

(i) Mortality. First Modification of the Equations of Part I 67
(ii) Second Modification of the Equations 69

SECTION XIII

(i) Discussion 70

Am. J. Hyg. 3 ('23) Jan. Suppl^{t}
31 Isoclines if $f = x^n$?

PART I

SECTION I

Introduction.—It is somewhat surprising that so little mathematical work should have been done on the subject of epidemics, and, indeed, on the distribution of diseases in general. Not only is the theme of immediate importance to humanity, but it is one which is fundamentally connected with numbers, while vast masses of statistics have long been awaiting proper examination. But, more than this, many and indeed the principal problems of epidemiology on which preventive measures largely depend, such as the rate of infection, the frequency of outbreaks, and the loss of immunity, can scarcely ever be resolved by any other methods than those of analysis. For example, infectious diseases may perhaps be classified in three groups: (1) diseases such as leprosy, tuberculosis, and (?) cancer, which fluctuate comparatively little from month to month, though they may slowly increase or decrease in the course of years; (2) diseases such as measles, scarlatina, malaria, and dysentery, which, though constantly present in many countries, flare up in epidemics at frequent intervals; and (3) diseases such as plague or cholera, which disappear entirely after periods of acute epidemicity.

To what are these differences due? Why, indeed, should epidemics occur at all, and why should not all infectious diseases belong to the first group and always remain at an almost flat rate? Behind these phenomena there must be causes which are of profound importance to mankind and which probably can be ascertained only by those principles of careful computation which have yielded such brilliant results in astronomy, physics, and mechanics. Are the epidemics in the second class of diseases due (1) to a sudden and simultaneous increase of infectivity in the causative agents living in affected persons; or (2) to changes of environment which favour their dissemination from person to person; or (3) merely to the increase of susceptible material in a locality due to the gradual loss of acquired immunity in the population there: or to similar or other causes? And why should diseases of the third class disappear, as they undoubtedly do, and diseases of the first class remain so persistently?—all questions which immediately and obviously present themselves for examination.

The whole subject is capable of study by two distinct methods which are used in other branches of science, which are complementary of each other, and which should converge towards the same results —the *a posteriori* and the *a priori* methods. In the former we commence with observed statistics, endeavour to fit analytical laws to them, and so work backwards to the underlying cause (as done in much statistical work of the day): and in the latter we assume a knowledge of the causes, construct our differential equations on that supposition, follow up the logical consequences, and finally test the calculated results by comparing them with the observed statistics.

Apparently the first *a posteriori* work of importance on epidemics was a communication by Dr. William Farr in 1866 (1), in which he maintained, in connection with an epidemic of cattle plague, that the course of the epidemic would follow a curve of which the third difference of the logarithm was a negative constant. In 1873–74, Dr. G. H. Evans (2) endeavoured to extend this law to other epidemics but ended by expressing disappointment with his results. More recently, however, Dr. J. Brownlee has continued the work in a series of papers (3–8). In the first of these he said that Dr. Farr's curve is the normal curve of probability, but, after studying a number of epidemics, found that Pearson's Type IV fitted better, and he concluded (p. 567) that "an epidemic is an organic phenomenon, the course of which seems to depend on the acquisition by an organism of a high grade of infectivity at the point where the epidemic starts, this infectivity being lost from that period till the end of the epidemic at a rate approaching to the terms of a geometrical progression." He admitted, however (p. 517), that "other factors, which are not clear, seem to come into play": but added "that the epidemic ends because of the lack of susceptible persons has no evidence in its favour, either from the form of the curve or from the facts" (see also pp. 500 and 501). It is obvious from the mere examination of many curves of epidemics that they are often remarkably symmetrical bell-shaped curves, which, however, frequently tend to fall somewhat more slowly than they rose; and in a later paper (4) Brownlee emphasises this feature of symmetry, and says (p. 2) that "the deduction from this phenomenon is direct and complete, namely, that the want of persons liable to infection is not the cause of the decay of the epidemic. On no law of infection which I have been able to devise would such a cause permit of epidemic symmetry. The fall must in all such cases be much more rapid than the rise, though, on the contrary, when asymmetry is markedly present the opposite holds." In still later papers (6, 7) he gives much evidence to show that measles and smallpox have distinct periods of recrudescence.

So far as I can ascertain, *a priori* researches on epidemiology were first commenced by myself since 1899 in connection with malaria. In 1904 I read a paper (9) concerned with the random migration of mosquitoes, a subject of vital importance in the theory of malaria which was subsequently dealt with, at my suggestion, by Prof. K. Pearson, whose researches were then employed by Brownlee (5). Subsequently I endeavoured to find the malaria equations by *a priori* reasoning (10), and in the second edition of my book (11), extended this method to a preliminary general "Theory of Happenings," employing chiefly the Finite Calculus (which is useful for malaria). Quite recently I published a very brief note (13), recording further advances; and Brownlee added another (7), showing how he arrived at some of his results.

The present paper is a much more advanced and general development of the work commenced in Section 66 (14) of the second edition of my book (11, p. 676). It deals merely with the curves which are theoretically obtained when we suppose that the infectivity-ratio remains constant or proportional to the number of individuals already affected, while, simultaneously, some of these are constantly "reverting," or losing immunity, and while both the affected and the unaffected groups are subject to special rates of birth, death, immigration and emigration—as defined in the following section. This thesis does not consider progressive decline of the infectivity; yet we shall see in Section VII that, contrary to what Dr. Brownlee supposed, it does often yield curves which are bell-shaped, nearly symmetrical, or, when otherwise, decline more slowly than they rose, that is, yields curves of the type frequently found in epidemics and does so without demanding any other cause for the decline than the exhaustion of susceptible material.

The thesis is stated in terms which are certainly academical, but into which it has been resolved only after very careful thought. It has not only been dealt with in considerable detail, but some of the more immediate consequences have been followed out—because I think that it must first be examined and used as a standard before we can proceed to discuss variations due to changes in the infectivity. As will be seen, the mathematics presents few difficulties; but, owing to the nature of the subject, I have thought it wise to labour the proofs somewhat more than some readers will require.

Beyond this point I have not endeavoured to go, because the application of the equations to numbers of known epidemics (which will ultimately be required) can be made only at great length. The paper is therefore purely theoretical; but it is one which I think will be needed for future studies.

Dr. Farr's communication (1) is not republished in his collected works ('Vital Statistics,' edited by N. A. Humphreys, London, 1885), but has recently been recovered with some difficulty by Dr. Brownlee. It is interesting to note that he seems to have attributed the decline of epidemics partly to the attenuation of infectivity, but also partly to "the fact that the individuals left are less susceptible of attack, either by the constitution or hygienic conditions, than those destroyed." The descriptions of his methods furnished by himself and by Evans are, however, almost unintelligible; but it is evident from "the calculated series by law" which he gives, that the third difference (not the second, as Brownlee said) of the logarithms of the series is a negative constant.

Some of Brownlee's equations are also difficult to interpret in terms of the numbers of individuals affected. I have therefore thought it best to proceed in this paper entirely on the basis of my own previous work.

SECTION II

The problem before us is as follows. Suppose that we have a population of living things numbering P individuals, of whom a number Z are affected by *something* (such as a disease), and the remainder A are not so affected: suppose that a proportion $h \, . \, dt$ of the non-affected become affected in every element of time dt, and that, conversely, a proportion $r \, . \, dt$ of the affected become unaffected, that is, revert in every element of time to the non-affected group: and lastly, suppose that both groups, the affected and the non-affected, are subject also to possibly different birth-rates, death-rates, and immigration and emigration rates in an element of time; then what will be the number of affected individuals, of new cases, and of the total population living at any time t?

For the solution of this and the subsidiary problems I have ventured to suggest the name "Theory of Happenings." It covers many cases which occur not only in pathometry but in the analysis of questions connected with statistics, demography, public health, the theory of evolution, and even commerce, politics, and statesmanship. The name *pathometry* (*pathos*, a happening) was previously suggested by myself in antithesis to *nosometry* (*nosos*, a disease) for the quantitative study of parasitic invasions in the individual.

SECTION III

(i) Let *ndt, mdt, idt, edt* denote respectively the nativity, mortality, immigration, and emigration rates of the non-affected part of the population in the element of time dt: and N*dt*, M*dt*, I*dt*, E*dt* denote the similar rates among the affected part. Then, as argued in my previous writings and as will be easily seen, the problem before us may be put in the form of the following system of differential equations:—

$\dot{P} = vA + VZ$

$$dP = (n - m + i - e)dt \, . \, A + (N - M + I - E)dt \, . \, Z. \qquad (1)$$

$$\begin{cases} dA = (n - m + i - e - h)dt \, . \, A + (N + r)dt \, . \, Z, & (2) \\ dZ = hdt \, . \, A + (-M + I - E - r)dt \, . \, Z. & (3) \end{cases}$$

From
A = P–
hA
rZ
nA
mA
iA
eA

Here $d\mathrm{P}$ consists only of the *variation-elements* n, m, i, e, N, M, I, E, with their proper signs, while $d\mathrm{A}$ and $d\mathrm{Z}$ contain also the *happening-element* h and the *reversion-element* r. Obviously $d\mathrm{P} = d\mathrm{A} + d\mathrm{Z}$, as it should do, the elements h and r disappearing in the summation because they denote only change of condition and not of being. The number of the non-affected who become affected in the element of time is $hdt \,.\, \mathrm{A}$, and, of course, this number pass over from group A to group Z—as shown in equations 2 and 3. Similarly $rdt \,.\, \mathrm{Z}$ is the number of the affected who revert in the element of time and therefore pass over from group Z to group A. The equations are not quite symmetrical, since N, which should appear in equation 3, appears in equation 2—because the progeny of the affected group, namely, $\mathrm{N}dt \,.\, \mathrm{Z}$, will generally be born not affected—an important fact which modifies almost all the results.

(ii) The variation-elements n, m, i, e, and N, M, I, E, may sometimes be functions of time, especially if the considered events extend over long periods: but it will quite suffice at present to take them as being constants. If we have sufficient data regarding them, we can generally calculate them by the methods of Section VIII for whatever small unit of time we adopt: otherwise they must be conjectured or assumed to begin with, and be then ascertained by a comparison of the integrated equations with known facts. It is convenient to write

$$\begin{aligned} v &\equiv n - m + i - e, \\ \mathrm{V} &\equiv \mathrm{N} - \mathrm{M} + \mathrm{I} - \mathrm{E}. \end{aligned} \tag{4}$$

In some kinds of happening which have no marked effect on the birth-rates, death-rates, and immigrations—such for instance as mild maladies, enlistments in civil professions, conversions to religious or political parties, entries into unlimited societies or trades-unions, etc.—we may have $n = \mathrm{N}$, $m = \mathrm{M}$, $i = \mathrm{I}$, $e = \mathrm{E}$, and consequently, $v \overset{4.13}{=} \mathrm{V}$. In other cases, quite possibly we may still have $v = \mathrm{V}$, though the individual items may be different. Generally in marriages $\mathrm{N} > n$; in accidents $\mathrm{M} > m$; in vaccinations $\mathrm{M} < m$; in total-abstainers $\mathrm{N} > n$ and $\mathrm{M} < m$; in military enlistments and in many diseases $\mathrm{N} < n$ and $\mathrm{M} > m$; while in certain alarming epidemics, especially cholera, plague, and malaria, we find in addition that

I < *i* and E > *e*—so that in these cases, which are particularly our present subject of study, we generally have *r* > F. In fatal accidents M = 1 and N, I, and E all = 0. When considering happenings among the *same* individuals, we also put N, I, and E at zero, though M may be anything; and in other special cases we omit various elements. If the surrounding population is not affected, I = 0; and if the affected individuals cannot move, E = 0: so that it will be seen that the equations can be made to cover 4.13 a wide field. The theory will sometimes also apply to inanimate objects—as, for instance, in commerce, where the variation-elements may be taken to mean manufacture, waste, exportation, and importation, though the original equations may have to be somewhat modified.

(iii) The *reversion-element* r may also at present be taken as a constant and be calculated for the same unit of time as is used for the other elements. In the case of independent happenings (to be defined presently) *rdt* means merely the proportion of affected individuals which become, in element of time, capable of being re-affected—as by divorce in marriage, and by discharge in certain employments. In dependent happenings, however, it implies also the loss of capacity for affecting others. Thus, in infectious diseases (which are dependent happenings), it implies loss both of immunity and of infectivity—not recovery from sickness, which is merely an episode of affectedness from disease. In some diseases, such probably as leprosy and organic diseases, *r* is zero or nearly zero; in others with long-continued immunity, such as many zymotic diseases, it is low; and in others again, with comparatively quick loss of immunity and infectivity, such possibly as nasal catarrh or dengue, it must be larger. In many diseases, however, it is quite unknown at present and must in fact be calculated from the integrated equations (which, it is hoped, will prove of use for this very purpose). In the case of slight accidents it is unity; in fatal accidents it is zero; and in snake-bite or heat-stroke, where recovery is quick if it occurs, it is high, though at the same time the death-rate M is also high. In controversial parties due to rational divergence of opinion it should be high; in ordinary party-politics, it is, in fact, very low.

(iv) The most important element is the *happening-element*, h. We should clearly understand that in most cases the happening, whatever it is, does not select only the non-affected, but tends to fall on both groups alike. If, however, it chances to fall upon individuals who are already affected, it merely re-affects them and does not cause them to pass from one group to another. Really the total number of individuals to whom the happening occurs in element of time is $hdt\,.\,\mathrm{P}$, that is, $hdt\,.\,(\mathrm{A}+\mathrm{Z})$. But the individuals numbering $hdt\,.\,\mathrm{Z}$ do not count because they are already affected, and therefore do not appear in the equations. On the other hand, the actual number of *new cases* (which we may denote by $\mathrm{F}dt$) is $hdt\,.\,\mathrm{A}$: and this quantity does appear in the equations 2 and 3. Or we may obtain F independently from the proportion

$$\frac{\mathrm{F}dt}{hdt\,.\,\mathrm{P}} = \frac{\mathrm{A}}{\mathrm{P}}. \tag{5}$$

so that, again, $\mathrm{F} = h\mathrm{A} = h\,(\mathrm{P} - \mathrm{Z})$. This is an important sub-function in all cases.

(v) Different kinds of happenings may be separated into two classes, namely, (*a*) those in which the frequency of the happening is *independent* of the number of individuals already affected ; and (*b*) those in which the frequency of the happening *depends* upon this quantity. To class (*a*) belong such happenings as many kinds of accidents and non-infectious diseases due to causes which operate, so to speak, from outside : and to class (*b*) belong infectious diseases, membership of societies and sects with propagandas, trade-unions, political parties, etc., due to propagation from within, that is, from individual to individual. In the former case, h or F will be constants ; in the latter case they will be functions of Z.

To proceed now to the integration of equations 1, 2 and 3—we may observe that the solutions, though they belong to the same class of functions, differ specifically in different cases. The integrals are easy to obtain by ordinary methods and are familiar in connection with many statistical, chemical and physical problems ; but the most elegant and useful method is to put all the cases into similar forms which can be brought by substitutions into the same immediately-integrable differential equation of the simplest type. We will first take a rapid survey of each solution as it is obtained.

SECTION IV

(i) *Independent Happenings.*—Here h or F is supposed to be constant. If h is constant, the happening falls on the same proportion, hdt, of the population in every element of time.

Putting P — Z for A, using v and V for the $\overset{\text{consc}}{\text{variation}}$-elements as in 4, and setting $x \equiv Z/P$, equations 1 and 3 become

$$\overset{= v(P-Z)+VZ}{dP/dt = vP - (v - V)\,xP}, \quad \text{incap. of} \tag{6}$$

$$\overset{= h(P-xP)+(V-N-r)xP}{dxP/dt = hP\,(1 - x) + (V - N - r)\,xP}. \tag{7}$$

$$= F + ..$$

mech?

and, of course,

$$= hP + (V - r)\,Z \quad p. 10$$

$$dxP/dt = xdP/dt + Pdx/dt.$$

Eliminating dxP/dt and dP/dt from these, we find that P cancels out also, so that we have $h\cancel{P}(1-x) + (V-N-r)x\cancel{P} = x(v\cancel{P} - \overline{v-V}\,x\cancel{P}) + \cancel{P}\dot{x}$

$$dx/dt = h - (h + v - V + N + r)\,x + (v - V)\,x^2. \tag{8}$$

This has one integral form if $\overset{4.10}{v = V}$, and another if $v \overset{p.20}{\neq} V$.

(ii) *The Equivariant Case.*—Here the happening is such that it does not affect the *sum* of the variation elements of the affected group—though, as already mentioned, the separate variation-elements may be different; that is, the total population is not altered. An example is, conversion to some philosophical creed without propagandism. Another example is, the happening of slight accidents—in which, however, r is generally unity; and a third example is the happening of a certain standard of wealth, which, we assume, ! tends to diminish the birth-rate, death-rate and immigration and emigration of the affected, each by an equal decrement, simultaneously. In this last case, however, as the children of the affected will tend to be born affected, we should remove N from equation 2 to equation 3 and omit it from equations 7 and 8. In all these cases $v = V$.

(iii) To integrate equation $\overset{\text{when } v = V}{8}$, however, write it in the form

$$dx/dt = \mathrm{K}(\mathrm{L} - x) \tag{9}$$

where $\mathrm{K} \equiv h + \mathrm{N} + r$ and $\overset{\text{p. 17}}{\mathrm{L}} \equiv h/\cancel{k}$. [Observe that the tangential, dx/dt, vanishes when $x = \mathrm{L}$ and becomes negative when $x > \mathrm{L}$.] Put $\underline{y \equiv \mathrm{L} - x}$; so that $x = \mathrm{L} - y$. Then, if L const (p. 13) N r const.

$= \frac{h}{h+N+r} - \frac{Z}{P}$ const.

$$dx/dt = -dy/dt = \mathrm{K}y,$$

$$1/y \,.\, dy = -\mathrm{K}dt. \tag{10}$$

Integrating both sides of this,

$$\log y = -\mathrm{K}t + \text{constant}.$$

If y_0 is the value of y at the beginning of the happening,

$$\log y_0 = \text{constant},$$

and $$y = y_0 \varepsilon^{-\mathrm{K}t}. \quad 4\,(30) \tag{11}$$

Therefore $$(\mathrm{L} - x) = (\mathrm{L} - x_0)\varepsilon^{-\mathrm{K}t},$$

and $$x = \mathrm{L} - (\mathrm{L} - x_0)\varepsilon^{-\mathrm{K}t}. \tag{12}$$

t

This gives the proportion, x, of the total population P who are affected at the time t, this proportion being x_0 when $t = 0$.

To find P, put $v - \mathrm{V} = 0$ in equation 6, and we obtain at once a differential equation of the same form as equation 9: so that

$$\mathrm{P} = \mathrm{P}_0 \varepsilon^{vt}, \tag{13}$$

where P_0 is the total population at the beginning of the happening, when $t = 0$. This is an important function which expresses the *natural* increase of the whole population due only to the natural variation-elements, v $\equiv n - m + i - e$. Of course v is small when the increase of population is not very rapid. When $\overset{= V}{v} = 0$, the population remains constant, the births, deaths, immigrations and emigrations annulling each other.

To find Z we have $\mathrm{Z} = x\mathrm{P}$.

In $\overset{\text{p. 12}}{\mathrm{III}}$ (iv), we defined F to be the number of $\overset{\text{when } I = 0}{\text{new cases}}$ and showed that $\mathrm{F} = h\,(\mathrm{P} - \mathrm{Z})$ $= hA$. If $\underline{f \equiv \mathrm{F/P}}$, we have, when $x_0 = 0$,

$f\mathrm{P} = h(\mathrm{P} - x\mathrm{P})$

$$f = h\,(1 - x) = h\,(1 - \mathrm{L} + \mathrm{L}\varepsilon^{-\mathrm{K}t}). = l + hLe^{-Kt} \tag{14}$$

$\dot f = h(-KLe^{-Kt})$

$\ddot f = hK^2Le^{-Kt}$

The functions for $\overset{=xP}{Z}$ and x ⑫ give what may be called respectively the *actual* and *proportional curves of affected individuals*: and those for $\overset{=fP}{F}$ and f ⑰ give respectively the *actual* and *proportional curves of new cases*.

To find t for any assigned value of x, we have

$$Kt = \log_e (L - x_0) - \log_e (L - x)$$
$$= 2{\cdot}3025851 \log_{10} [(L - x_0)/(L - x)]. \quad (15)$$

The following equations are required for ~~analysing~~ describing the curves x and f:— Fuss

$$= \tfrac{d}{dt}(-\dot{y}) = \tfrac{d}{dt}(Ky) = K\dot{y} \; [\text{if } \overset{= \frac{h}{L} \; N+r}{K} \text{ const.}] = K(-Ky)$$

$$d^2x/dt^2 = - K^2 (L - x). \quad (16)$$

⑭ $= h(-\dot{x}) \; [\text{if } h \text{ const.}] = h(Ky)$

$$df/dt = - hK (L - x). \quad (17)$$

$$\overset{= -hK(-\dot{x})}{d^2f/dt^2} = hK^2 (L - x). \quad (18)$$

(iv) In order to analyse the curve ⑫ x (of affected individuals) we first observe that as h, N, r are always positive, K and L are also always positive and L is always less than unity, while ε^{-Kt} diminishes and finally vanishes! as t increases. Hence when t is very large, x reaches the limit L (equation 12) and never exceeds it if $x_0 \leqq L. < 1$ Its tangential (equation 9) No! begins at the value KL ($= h$) when $x_0 = 0$, always remains positive, and gradually diminishes to zero when $x = L$ ⑫ ~~and $t = \infty$~~. Its second tangential (equation 16) is always negative. Hence as t increases from zero to infinity, x also increases, but with a diminishing increment, until it approximates to the asymptote L; and it is always con~~vex~~ cave to the axis of t. $\dot{x} = K$ cf. Melion

On the other hand the curve f (equation 14, for w.h. $x_0 = 0$ of new cases) begins at its greatest value h (~~when~~ $x_0 = 0$) and constantly diminishes, its tangential (equation 17) being always negative. Its second tangential is always positive, and the curve for f is, therefore, convex to the axis of t, and when t is large, Not from ⑱ f approaches its final value h (1 — L), $= L(h + N$ that is (N + r)L, which is always whatever x_0 ? less than h. Let this constant be denoted by $l. = h(1-L)$

It may have been thought that, at least in the case of constant i.e. indep. happenings, the whole population would become affected; but it has been shown that this is not the case, because x cannot exceed

L. which is a proper fraction. What is it then that limits the value of x? Equation 8 may be written (when $v = \mathrm{V}$)

$$dx/dt = h\,(1 - x) - (\mathrm{N} + r)\,x = f - (\mathrm{N} + r)\,x, \tag{19}$$

and dx/dt vanishes when $f = (\mathrm{N} + r)\,x$. When it vanishes, ⑨ $x = \overset{\text{whatever } h}{\mathrm{L}}$, ∴ $t \to$ and ⑭ $f = l$; that is, the proportion of affected individuals comes to a limit when the proportion of new cases exactly balances the proportion of recoveries and births (it being assumed by supposition that the progeny of the affected are born not affected). If, however, N and r are zero (and only then), $\mathrm{L} = 1$, and the whole population does become affected in time. It should also be observed that ⑭ $\frac{l}{h} = \cancel{h}\,(1 - \mathrm{L}) = \cancel{h}\mathrm{A_L}/\mathrm{P}$, where $\mathrm{A_L}$ is the final number $\overset{\text{whatever } h}{\text{of}}$ non-affected individuals when $x = \mathrm{L}$ and t is very large; that is to say, even after the limit of x has been nearly approached, the happening continues and constantly affects the same proportion of the non-affected population as before—but now A also comes to a limit, namely $\mathrm{P}\,(1 - \mathrm{L})$.

! The easiest $\overset{\text{clumsy}}{\text{way}}$ to assess the increase of x as a function ⑫ of t is to give to $\varepsilon^{\mathrm{K}t}$ the successive values 1, 2, 3, 4 … 10, 100, 1000 …. Then it is clear from equation ⑫ that, if $x_0 = 0$, x will have the corresponding values 0, $\tfrac{1}{2}\mathrm{L}$, $\tfrac{2}{3}\mathrm{L}$, $\tfrac{3}{4}\mathrm{L}$, … $0\cdot9\mathrm{L}$, $0\cdot99\mathrm{L}$, $0\cdot999\mathrm{L}$, … and t will have the values

$1 - e^{-kt}$

$$\frac{2\cdot302\ldots}{\mathrm{K}}\log_{10}(1,\ 2,\ 3,\ \ldots\ 10,\ 100,\ 1000\ \ldots).$$

An important point is reached when $\varepsilon^{\mathrm{K}t} = 10$, for then

$kt = \ln 10 \doteq 2{\cdot}3$

$$t = \frac{1}{\mathrm{K}}\,2\cdot302\ldots \equiv \tau \text{ (say)}. \tag{20}$$

At this point x has increased to 9/10ths of its ultimate value L—that is, most of the change of which is it capable has been effected, and it increases only slightly afterwards, namely only by $0\cdot09$ and $0\cdot009$, when $t = 2\tau$, and $= 3\tau$, and so on. On the other hand, x reaches half its ultimate value very quickly when $t = \frac{1}{3}\tau$ nearly.

$\log 2 = {\cdot}3010$

However large t may be, $\varepsilon^{-\mathrm{K}t}$ never quite vanishes, and therefore x never quite reaches L. It is convenient, then, to find the value of

t when the number of affected individuals reaches the limit *less one individual*—that is, when $Z = LP - 1$. This is obtained at once from equation 15, since $L/(L - x) = LP/(LP - Z)$; and the required value is

$$t = \tau \log_{10} LP \equiv \tau' \text{ (say)}. \qquad (21)$$

For example, if $L = \frac{1}{2}$ and $P = 2{,}000{,}000$ and remains nearly constant, the number of affected individuals is 900,000 when $t = \tau$, and is 999,999 when $t = 6\tau$.

[The ratio x/t is the tangent of slope of the radius-vector from the origin to the point (x, t).] If $x_0 = 0$ the value of this tangent is h, that is, dx_0/dt, when $t = 0$; and is $9L/10\tau$, that is, $0{\cdot}39086h$, when $t = \tau$; so that, as the curve of x increases from zero to nine-tenths of its ultimate value, it or the curve lies wholly within the angle formed by these vectors. And the magnitude of this angle depends only upon h, and not upon the other elements, N and r.

$x = 1 - e^{-kt}$

$\dot{x} = ke^{-kt}$

$= he^{-kt} . e^{(N+r)t}$

$\dot{x} = he$

4.

The constants L, l, and τ may be written in detail,

$$L = \frac{1}{1 + (N + r)\,h^{-1}}, \qquad l = \frac{1}{h^{-1} + (N + r)^{-1}} = (N+r)L = \frac{L}{(N+r)^{-1}},$$

$$\tau = \frac{2{\cdot}302\ldots}{h + N + r}. \qquad (22)$$

From these equations it will be easily seen that if h increases while N and r remain constant, L and l will increase and τ will diminish; but if h remains constant while N or r or both increase, L and τ will diminish and l will increase—just as may be expected. If $r = 1$ and $N \neq 0$, $l > L$, since $l = (N + r)\,L$; that is, as all the cases recover as soon as they occur, the total number of cases will not exceed the new cases during unit of time—as happens with slight accidents such as mosquito bites, where there is nothing in one happening to prevent a second or third occurrence of it. If $N + r = 0$, $L = 1$, $l = 0$, and $\tau = h$, and if $h = N + r$, $L = \frac{1}{2}$, $l = \frac{1}{2}h$, and, $\tau = 1{\cdot}151\ldots$

(v) It remains to examine some questions which require for their solution the integrals of P, Z, A, and F—all easily obtained.

From equation 13 we have at once

$$\int P\,dt = \int P_0 \varepsilon^{vt}\,dt = P/v + \text{constant},$$

$$\int_0^t P\,dt = (P - P_0)/v. \tag{23}$$

To find the integrals of Z, A, and F, we integrate both sides of equation 7 as there written, or as written in the form

$$dZ/dt = hP - (K - V)\,Z, \quad \left\{ qZ = e^{-\int(V-K)dt}\int(hP e^{\int(V-K)dt} \text{ not } c \right.$$

$$= F + \cancel{hZ} + V - h - (N - r)\,Z \quad = \int hP\,dt + Ce^{(K-V)c}$$

and obtain

$$Z = \underset{\text{New cases}}{\int F\,dt} + (V - N - r)\int Z\,dt, + Z_0 \tag{24}$$

or

$$= h\int P\,dt - (K - V)\int Z\,dt\,; + Z_0 \tag{25}$$

whence

$$+ Z_0 = h\,\frac{P-P_0}{v} - (K-V) \quad '' \quad + Z_0$$

$$\int_0^t Z\,dt = J\,\frac{P - P_0}{v} - J\,\frac{Z - Z_0}{h}, \tag{26}$$

$$\int_0^t A\,\overset{=\int(P-Z)dt}{dt} = (1 - J)\,\frac{P - P_0}{v} + J\,\frac{Z - Z_0}{h}, \tag{27}$$

where

$$J = h/(K - V) \quad \text{and} \therefore 1 - J = -\frac{V - N - r}{K - V}.$$

And, of course,

$$\int_0^t F\,dt = h\int_0^t A\,dt.$$

The same equations can be readily obtained by direct integration from equations $\overset{Z=hP}{12}$, $\overset{P}{13}$, etc.

The concrete interpretation of these definite integrals of P, Z, and A, is that they express the total number of *time-units lived* during the period t by the whole population, by the affected population, and by the non-affected population respectively.

For example, suppose that a population P_0 has doubled itself in 1000 time-units, then, ⑬ since $1000v = \log_e 2, \therefore v = 0 \cdot 00069315$, and the total number of time-units lived by all the individuals together is ㉓ $P_0(2 - 1)/v$—that is, $1442 \cdot 7\ P_0$ time-units.

It should be observed that if $\underline{v}$ is very small, each side of equation 23 takes an indeterminate form, the value of which is the value of $P_0\,(\varepsilon^{vt} - 1)/v$ when $v = 0$—namely, $\underline{tP_0.}$ That is, when the population remains nearly constant, the total number of time-units lived during t is tP_0, as we should expect.

The same integrals divided by t will give the *mean numbers* of the total, the affected, and the non-affected populations respectively during the period t. Thus, if the happening is such that it does not affect the normal rate of increase or decrease of population, then the mean population will be $(P - P_0)/(\overset{(3)}{\log_\varepsilon} P/P_0)$—as we should expect. 4.

When multiplied by the appropriate constants, these integrals will also give the total number of the *variation-events*, births, deaths, immigrations and emigrations. Thus the total *progeny* of the non-affected will number $n \int_0^t A\,dt$, and the total *deaths* among the affected will number $M \int_0^t Z\,dt$. The sum of the natural variation-events in the whole population is $(n - m + i - e) \int_0^t P\,dt$, that is, $P - P_0$. Similarly $r \int_0^t Z\,dt$ gives the total *reversions*: and the integrals multiplied by h give the total number of time-units lived in which the *happenings* have occurred—which is the same as the number of individuals to which the happenings have occurred. Thus $h \int_0^t A\,dt$ is the total number of happenings (not necessarily for the first time) among the non-affected—that is the total number of *new cases*, $\int_0^t F\,dt$. → p. 14.

It is now easy to interpret equations 24, 25, 26, 27. Since $V - N - r = I - E - M - r$, equation 24 means that Z, the total number of affected individuals remaining alive at the time t, is equal to the sum of all the new cases plus the affected immigrations and less the affected emigrations, deaths, and recoveries. Also as $K - V = h + r + M + E - I$, equation 25 means that the same quantity Z is equal to the total happenings plus the affected immigrations, less the affected emigrations, deaths, and recoveries, ?

and also less the happenings among the affected—which, as stated in III (iv), do not count.

Section V

(i) *Independent Happenings: the General Case:* $v \overset{p.13}{\neq} V$.—The method of working employed for the equivariant case has been given at some length because it will serve also for the other cases. We now proceed to integrate the general equation 8, namely,

$$dx/dt = h - (h + v - V + N + r)x + (v - V)\,x^2.$$

Write this in the form

$$dx/dt = K\,(L - x)\,(L' - x), \tag{28}$$

where $K = v - V$ and L and L′ are the two roots of the quadratic in x. Let $L = \alpha - \beta$ and $L' = \alpha + \beta$, where $\alpha = (h + N + r + K)/2K$ and $\beta^2 = \alpha^2 - h/K$. These roots are always real because

$$(h + N + r + K)^2 - 4hK$$
$$= (h - K)^2 + 2(h + K)(N + r) + (N + r)^2,$$

of which the left side is positive when K is negative, and the right side is positive when K is positive. And the roots are both positive when $v > V$. Observe that dx/dt vanishes when $x = L$ or L′, if this ever occurs.

(ii) Put $\frac{dy}{dx} = \frac{-1}{L-x} + \frac{L'-x}{(L-x)^2} = \frac{-(L-x)+L'-x}{(L-x)^2} = K(L-x)(L'-x)\,\frac{L'-L}{(L-x)^2}$

$$y \equiv (L' - x)/(L - x), \quad x = L - (L' - L)/(y - 1). \tag{29}$$

Then $= K\,\frac{L'-x}{L-x}\,y = (L'-L)\,\frac{L-x}{L-L'} \propto x$

$$\frac{dx}{dt} = \frac{L' - L}{(y - 1)^2}\,\frac{dy}{dt} = Ky\,(L - x)^2 = Ky\left(\frac{L - L'}{y - 1}\right)^2;$$

$$dy/dt = K\,(L' - L)\,y, \qquad y = y_0\varepsilon^{K\,(L'-L)\,t}, \tag{30}$$

as in equation $\overset{p.16}{11}$. Therefore, as $L' - L = 2\beta$,

$$\frac{L' - x}{L - x} = \frac{L' - x_0}{L - x_0}\,\varepsilon^{2K\beta t}\,; \therefore \frac{\overset{2\beta}{L'-L}}{L-x} = \frac{(L'-x_0)\varepsilon^{2K\beta t} - (L-x)}{L-x_0} \tag{31}$$

or

$$x = L - \frac{2\beta\,(L - x_0)}{(L' - x_0)\varepsilon^{2K\beta t} - (L - x_0)}$$

$$= L\left\{1 - \frac{L' - \overset{i.e.\ 2\beta}{L}}{L'\varepsilon^{2K\beta t} - L}\right\} \quad \text{when} \quad x_0 = 0. \tag{32}$$

Acc. No. 17337.

The first tangential of this is given in equation 28; the second tangential is

$= k(-L + x - L' + x)\dot{x} = K(-2\alpha + 2x)\dot{x}(L-x)(L'-x)$

$$d^2x/dt^2 = -2K^2(\alpha - x)(L - x)(L' - x). \tag{33}$$

(iii) The analysis of x, *the proportional curve of affected individuals*, depends upon whether K, that is $v - V$, is positive or negative. If it is positive, we are generally concerned with cases of Injurious Happenings; because, as explained in Section III (ii), the nativity, mortality, immigration and emigration are likely to be respectively less, greater, less, and greater among the affected in such cases than among the non-affected. Non-infective diseases should be examples in point. On the other hand, if $v - V$ is negative, we shall be generally concerned with cases of Beneficial Happenings, which improve the natural variation elements of the affected—such for example as conversions to total abstinence are said to be (if they do not diminish i and increase e in places where they prevail, such as the Prohibition States in America).

If K is positive, $\varepsilon^{2K\beta t}$ in equations 31 and 32 increases indefinitely with t (since β is the positive value of the radicle in the roots L and L′); so that x approaches the limit L, that is, $\alpha - \beta$; and never exceeds it, and therefore never reaches the greater root L′, that is, $\alpha + \beta$. Therefore dx/dt is always positive for values of x between zero and L; and d^2x/dt^2 is always negative for the same values, since $\alpha > \alpha - \beta$. Hence in this case x has a form similar to that which it has in the equivariant case. That is, it begins (when $x_0 = 0$) at zero at an angle of which the tangent is h; constantly increases with a decreasing increment, so that it always remains concave to the axis of t; and finally approaches the limit L when t is very large.

$v >$

When K is negative, the quadratic expression in equations 8 and 28 has only one change of sign, and therefore only one positive root, L′ or $\alpha + \beta$. The other root, L, does not concern us because we consider only positive values of x. In this case, $\varepsilon^{2K\beta t}$ diminishes indefinitely as t increases, and x finally reaches the limit L′. Here, too, dx/dt always remains positive for possible values of x, that is, while t varies from zero to positive infinity; so that x again always increases from zero to L′. If $h + v - V + N + r$, that is, $2K\alpha$,

$v <$

Descart

$\therefore \alpha < 0$

is positive, d^2x/dt^2 is always negative in this case also, and x is therefore concave to the axis of x as before. But if this expression is negative, d^2x/dt^2 will be positive at first, but will change sign when $x = \alpha$, and finally vanish when $x = \mathrm{L}'$; so that, in this case, x will be first convex and then concave to the axis of t. This is the only difference in general form. It is seldom that $\mathrm{V} - v$ can be greater than $h + \mathrm{N} + r$, which is always positive; hence x has the same form in most cases where h is a constant.

(iv) For the proportional curve of new cases, f, we have

$$f = h\,(1 - x)\,; \tag{34}$$

$$df/dt = -\,h\,dx/dt, \qquad d^2f/dt^2 = -\,h\,d^2x/dt^2. \tag{35}$$

Hence, when $x_0 = 0$, f begins at the value h and, as t increases, gradually falls to the limit $h\,(1 - \mathrm{L})$ when K is positive, and to the limit $h\,(1 - \mathrm{L}')$ when K is negative. The curvature is the opposite of that of x—as in the equivariant case.

To find the value of P, the total population, write equation 6 in the form,

$$1/\mathrm{P}\,.\,d\mathrm{P}/dt = v - \mathrm{K}x,$$

Therefore

$$\begin{aligned}\log \mathrm{P} &= vt - \int \mathrm{K}x\,dt \\ &= vt - \mathrm{KL}t + \int \frac{2\mathrm{K}\beta\varepsilon^{-2\mathrm{K}\beta t}}{y_0 - \varepsilon^{-2\mathrm{K}\beta t}}\,dt \\ &= (v - \mathrm{KL})\,t + \log\,(y_0 - \varepsilon^{-2\mathrm{K}\beta t}) + \text{constant.}\end{aligned}$$

On evaluating the constant we find that we can arrange P in either of two ways, namely,

$$\mathrm{P} = \mathrm{P}_0\varepsilon^{vt}\frac{y_0 - \varepsilon^{-2\mathrm{K}\beta t}}{y_0 - 1}\,\varepsilon^{-\mathrm{KL}t} \quad \text{or} \quad \mathrm{P} = \mathrm{P}_0\varepsilon^{vt}\frac{y_0\varepsilon^{2\mathrm{K}\beta t} - 1}{y_0 - 1}\,\varepsilon^{-\mathrm{KL}'t}. \tag{36}$$

If $x_0 = 0$ and t is large, these become, when $\mathrm{K} > 0$ and $\mathrm{K} < 0$, respectively,

$$\mathrm{P} = \mathrm{P}_0\varepsilon^{vt}\,\frac{\mathrm{L}'}{\mathrm{L}' - \mathrm{L}}\,\varepsilon^{-\mathrm{KL}t} \quad \text{or} \quad \mathrm{P} = \mathrm{P}_0\varepsilon^{vt}\,\frac{-\,\mathrm{L}}{\mathrm{L}' - \mathrm{L}}\,\varepsilon^{-\mathrm{KL}'t}. \tag{37}$$

The former expressions are suitable when K is positive, and the latter when it is negative. Thus in the case of injurious happenings when K is positive, the population will diminish indefinitely unless the natural increase denoted by v is large enough to compensate for the decrease, KL, due to the happening (L being positive in this case). In the case of beneficial happenings when K is negative, L is also negative while L′ remains positive, so that the happening enhances the natural increase of the population due to the happening.

As with equivariant happenings, x never quite reaches either L or L′; and it will therefore be useful to find the value of t when the affected population equals the limit less one individual, that is, when $Z = LP - 1$ or $L'P - 1$, or when $x = L - 1/P$, or $L' - 1/P$. This can easily be ascertained from equations 29 and 30, and we have to find t from the equation $1/P = (L' - L) / (y - 1)$ when K is positive, and from the equation $1/P = y\,(L' - L) / (y - 1)$ when K is negative. From these we have when $x_0 = 0$

$$\begin{aligned} (KL' - v)\,t &= \log_\varepsilon LP_0 \qquad (K > 0), \\ (KL - v)\,t &= \log_\varepsilon L'P_0 \qquad (K < 0). \end{aligned} \tag{38}$$

To find the integral of P, we write P in the form (when $x_0 = 0$),

$$P = \frac{L'P_0}{L' - L}\,\varepsilon^{(v - KL)\,t} - \frac{LP_0}{L' - L}\,\varepsilon^{(v - KL')\,t}.$$

Therefore

$$\int_0^t P\,dt = P_0\varepsilon^{vt}\frac{L'}{L' - L}\,\frac{\varepsilon^{-KLt} - \varepsilon^{-vt}}{v - KL} - P_0\varepsilon^{vt}\frac{L}{L' - L}\,\frac{\varepsilon^{-KL't} - \varepsilon^{-vt}}{v - KL'}. \tag{39}$$

The integrals of Z and F are found from this by the same methods as were used in equations 24 and 25, or by integrating both sides of equation 6. From the latter we have

$$(v - V)\int_0^t Z\,dt = v\int_0^t P\,dt - (P - P_0). \tag{40}$$

The remarks made in Section IV (v) apply here also.

SECTION VI

Independent Happenings : Constant New Cases.—A third example of independent happenings remains to be considered, namely, the one referred to at the beginning of Section IV, in which F, the number of new cases, is always a constant. The only instances of this which I can recall are the cases of certain societies (such as the Royal Society) which elect a fixed number of new members every year. Indeed, this process of increment seems to be too artificial to be seen in nature, unless, perhaps, in certain cases of seeding or spore formation, and we need only note the form of the principal function for future reference if required.

Here, since F is a constant, let $F = c$. Then equations 6 and 7 become

$$dP/dt = vP - (v - V)\,Z, \tag{41}$$

$$\begin{aligned} dZ/dt &= c + (V - N - r)\,Z, \qquad (42)\\ &= K\,(L - Z). \end{aligned}$$

where $K \equiv M - I + E + r$ and $L \equiv c/K$. The solution of this is given in equation 12, except that Z is here substituted for x. If $Z_0 = 0$, we have

$$Z = c\,\frac{1 - \varepsilon^{-(M-I+E-r)t}}{M - I + E + r}.$$

It is to be hoped that the case of the Royal Society is at least an equivariant one.

SECTION VII

(i) *Dependent Happenings: Proportional Happening.*—Referring to the definition of dependent happenings in Section III (v), we see that in such h must be a function of Z, and consequently of t. First consider the case in which each affected individual affects or infects c other individuals in unit of time, c being a constant. This may be taken as being a first approximation to the study of those very important happenings, the infectious diseases.

The total number of the happenings which occur in element of time will be $cZdt$; but, as in the preceding cases, some of these may chance upon individuals who are already affected, and the number of new cases in element of time, namely, Fdt, will therefore most probably be given by the proportion

$$Fdt/cZdt = A/P, \tag{43}$$

that is $F = cZ\,(1 - x)$, and $h = cx$.

(ii) Equations 1 and 3 now become

$$dP/dt = vP - (v - V)\,xP, \tag{44}$$

$$dxP/dt = cxP\,(1 - x) + (V - N - r)\,xP. \tag{45}$$

Treating these as in Section IV (i), we find that, as before, on eliminating dxP/dt and dP/dt, P also cancels out, and we have

$$dx/dt = (c - v + V - N - r)\,x - (c - v + V)\,x^2, \tag{46}$$

$$= Kx\,(L - x), \tag{47}$$

where $K \equiv c - v + V$ and $L \equiv 1 - (N + r)/K$. Here again dx/dt vanishes when $x = L$. Integrating by substitution as before, put

$$y = (L - x)/x, \qquad x = L/(1 + y). \tag{48}$$

Hence

$$-\frac{L}{(1+y)^2}\frac{dy}{dt} = K\frac{L}{1+y}\left(L - \frac{L}{1+y}\right) = KL^2\frac{y}{(1+y)^2},$$

$$dy/y = -KL\,dt, \qquad y = y_0\varepsilon^{-KLt}, \tag{49}$$

$$x = \frac{L}{1 + (L/x_0 - 1)\,\varepsilon^{-KLt}}$$

$$= \frac{Lx_0\varepsilon^{KLt}}{x_0\varepsilon^{KLt} + (L - x_0)}. \tag{50}$$

Before analysing this it is advisable to obtain the integrated expressions for P and Z. Dividing equation (44) by P and integrating both sides, we have

$$\log P = vt - (v - V)\int x dt$$

$$= vt - (v - V)/K\,.\,\log\,(x_0\varepsilon^{KLt} + L - x_0) + \text{constant},$$

whence

$$P = P_0\varepsilon^{vt}\left(\frac{L}{x_0\varepsilon^{KLt} + L - x_0}\right)^{(v-V)/K} \tag{51}$$

and

$$Z = Z_0\varepsilon^{vt}\varepsilon^{KLt}\left(\frac{L}{x_0\varepsilon^{KLt} + L - x_0}\right)^{c/K}, \tag{52}$$

where the fraction within the brackets obviously $= (x/x_0)\,\varepsilon^{-KLt}$.

The explicit value of t is given by

$$\mathrm{KL}t = \log_\varepsilon \frac{\mathrm{L} - x_0}{x_0} - \log_\varepsilon \frac{\mathrm{L} - x}{x}. \tag{53}$$

(iii) For analysing the *proportional curve of affected individuals,* x, the first tangential is given in equation 47. The second and third tangentials are

$$d^2x/dt^2 = \mathrm{K}^2x\,(\mathrm{L} - x)\,(\mathrm{L} - 2x), \tag{54}$$

$$d^3x/dt^3 = \mathrm{K}^3x\,(\mathrm{L} - x)\,(\mathrm{L}^2 - 6\mathrm{L}x + 6x^2). \tag{55}$$

From these and from equations 50, it will be seen that all the tangentials of x are zero when $x = 0$ and $= \mathrm{L}$. If x_0 is very small, the curve begins at nearly zero when $t = 0$; and then increases very slowly at first and more rapidly afterwards, remaining convex to the axis of t to begin with. When, however, $x = \frac{1}{2}\mathrm{L}$, there is a change of curvature and x becomes concave to the axis of t, and then finally approaches the ultimate limit L when t is very large, and never exceeds that limit. Its maximum rate of increase, namely $\frac{1}{4}\mathrm{KL}^2$, is reached when $x = \frac{1}{2}\mathrm{L}$; and then

$$\mathrm{KL}t = \log\left[(\mathrm{L} - x_0)/x_0\right].$$

The curve is now seen to be a symmetrical one with this point as the centre of symmetry. Moving the origin to the point by putting $x = \frac{1}{2}\mathrm{L} + x'$ and $t = 1/\mathrm{KL} \,.\, \log\left[(\mathrm{L} - x_0)/x_0\right] + t'$, we obtain

$$x' = \tfrac{1}{2}\mathrm{L}(1 - \varepsilon^{-\mathrm{KL}t'})/(1 + \varepsilon^{-\mathrm{KL}t'}). \tag{56}$$

Substituting $-t'$ for t' in this, we find that x' also merely changes its sign without changing its numerical value. In fact, x has the general shape of a long-drawn-out letter S.

(iv) The curve dx/dt is also a symmetrical one with its centre of symmetry at the same distance from the primary origin, namely $\log\left[(\mathrm{L} - x_0)/x_0\right]$. When x_0 is small, the curve increases slowly at first, then rapidly; reaches a maximum value, $\frac{1}{4}\mathrm{KL}^2$, at the centre of symmetry, and then falls just as it rose. It has two changes of curvature which are the two roots of the quadratic expression in d^3x/dt^3, namely when $x = (\frac{1}{2} \mp \sqrt{\frac{1}{12}})\,\mathrm{L}$: that is when

$$x' = \mp 0{\cdot}28867\mathrm{L},$$

$$\text{and} \qquad \mathrm{KL}t' = \mp \log[(\sqrt{3} - 1)/(\sqrt{3} + 1)] = \mp 1{\cdot}31697. \tag{57}$$

At these points the ordinate of dx/dt, that is, of $\mathrm{K}x\,(\mathrm{L} - x)$, is

$$\mathrm{KL}^2\,(\tfrac{1}{2} + \sqrt{\tfrac{1}{12}})\,(1 - \tfrac{1}{2} - \sqrt{\tfrac{1}{12}}),$$

that is $\frac{2}{3}\,\frac{1}{4}\mathrm{KL}^2$, or two-thirds of the maximum ordinate at the centre of symmetry; and the ordinate is, of course, the same at both points of change of its curvature, as also at other points equidistant from the centre of symmetry.

The curve of dx/dt is therefore a *regular bell-shaped curve*, very similar to those often found in epidemics.

By the conditions of this kind of happening, x_0 and Z_0 can never be zero, for if they were there would be no new cases. It is convenient therefore to take Z_0 as being one individual—that is, $\mathrm{Z}_0 = 1$, $x_0 = 1/\mathrm{P}_0$, and $dx_0/dt = \mathrm{K}\,(\mathrm{LP}_0 - 1)/\mathrm{P}_0^2$, or nearly KL/P_0.

As with the kinds of happening previously considered (see equation 21), x and Z never quite reach the limits L and LP respectively; it is useful then to find when the number of affected individuals reaches the limit less one individual, that is when $\mathrm{Z} = \mathrm{LP} - 1$ and $x = \mathrm{L} - 1/\mathrm{P}$. If $\mathrm{Z}_0 = 1$ and $x_0 = 1/\mathrm{P}_0$ as just suggested, then, owing to the symmetry of the curves, this figure will be reached when $x = \mathrm{L} - x_0$; that is (equation 53) when

$$\begin{aligned} \mathrm{KL}t &= \log[(\mathrm{L} - x_0)/x_0] - \log[x_0/(\mathrm{L} - x_0)], \\ &= 2\log(\mathrm{LP}_0 - 1) = 2\mathrm{KL}\tau(\text{say}). \end{aligned} \tag{58}$$

Thus if the total population is large and L is not very small, τ, the abscissa of the centre of symmetry when $x_0 = 1/\mathrm{P}_0$, is very nearly equal to $1/\mathrm{KL}\,.\,\log \mathrm{LP}_0$; and at twice this period from the beginning the limit of affected individuals less one will be reached. We observe also that when $t = 2\tau$ the ordinate of dx/dt will be the same as it was at the beginning, namely, nearly KL/P_0.

If $x_0 = 1/P_0$, the tangent of the radius-vector from the origin to the summit of dx/dt is $\frac{1}{4}KL^2/\tau$, that is

$$K^2L^3/4 \log (LP_0 - 1). \tag{59}$$

From equations 48 and 49 we have

$$\varepsilon^{KLt} = \frac{x}{x_0}\frac{L - x_0}{L - x} = \frac{xP}{x_0P_0}\frac{LP_0 - x_0P_0}{LP - xP} = \frac{Z}{Z_0}\frac{LP_0 - Z_0}{LP - Z}. \tag{60}$$

This provides a useful formula for a cursory estimation of the increase of Z with respect to t when the latter is small. For if P_0 and P, Z_0 and Z are not very different, and LP_0 is large and Z_0 small—as at the commencement of the happening—then this equation is nearly

$$Z = Z_0\varepsilon^{KLt}.$$

Suppose that Z_0 is one individual ($t = 0$), then $Z = 2, 3, 4, \ldots$ individuals as ε^{KLt} increases through the values $2, 3, 4, \ldots$; that is, as t increases through the values $1/KL \,.\, \log_{10}$* $(2, 3, 4, \ldots)$. This same thing does not hold true, however, when Z is larger.

Similarly, we may have at first $F = cZ_0\varepsilon^{KLt} - \frac{1}{P_0} c\,(Z_0\varepsilon^{KLt})^2$.

(v) The *actual* and *proportional curves of new cases* are respectively :

$$F = cZ\,(1 - x),$$

$$f = cx\,(1 - x). \tag{61}$$

To analyse f we have

$$df/dt = cKx\,(L - x)\,(1 - 2x), \tag{62}$$

$$d^2f/dt^2 = cK^2x(L - x)\,\{L - 2\,(2L + 1)\,x + 6x^2\}. \tag{63}$$

And, since $L = 1 - (N + r)/K$, we may also write f in the form

$$Kf/c = dx/dt + (N + r)\,x. \tag{64}$$

The form of the curve f depends upon the value of L. If $L = 1$, that is, if $(N + r)/K$ is zero or very small, the curve f equals or approximates to the curve dx/dt multiplied by the constant c/K—

---

* As in original, but may be $\log_\epsilon$.

which is generally greater than unity. That is, in this case f is a symmetrical bell-shaped curve with the properties just described. In any case, however, f is such a curve plus the constantly-increasing function $c(\mathrm{N} + r)x/\mathrm{K}$, or $c(1 - \mathrm{L})x$. We may, therefore, expect that, at first, when x is small f will follow the graph of dx/dt, but that as x increases, the ordinates of f will become increasingly greater than those of dx/dt. At last, when t is very large and x approximates to L, dx/dt becomes nearly zero and f approaches the constant value $c\mathrm{L}(1 - \mathrm{L})$, at which it remains indefinitely. Thus f is quite symmetrical only if $\mathrm{L} = 1$; but is nearly symmetrical if L is only a little less than unity.

It is evident from equation 62 that df/dt vanishes when x equals either L or $\frac{1}{2}$. That is, if $\mathrm{L} > \frac{1}{2}$, f reaches a maximum when $x = \frac{1}{2}$, the ordinate of this maximum being exactly $c\frac{1}{2}(1 - \frac{1}{2})$, or $\frac{1}{4}c$; and as x increases above this value, f then decreases, and *falls* towards its ultimate constant value $c\mathrm{L}(1 - \mathrm{L})$ as x approaches L. If $\mathrm{L} = \frac{1}{2}$ exactly, df/dt vanishes only once (and that not quite), namely, when $x = \mathrm{L} = \frac{1}{2}$, so that f now *always increases* until it reaches its maximum, $\frac{1}{4}c$—which is also its ultimate value. If $\mathrm{L} < \frac{1}{2}$, however, x never attains the value $\frac{1}{2}$, and consequently df/dt has no vanishing point for any values of x between x_0 and L, and therefore again *always increases* for all considered values of x and therefore of t. Hence the curve of f has two forms: (Type I) if $\mathrm{L} > \frac{1}{2}$, an irregular bell-shaped form (becoming regular if $\mathrm{L} = 1$); and (Type II) if $\mathrm{L} \ngtr \frac{1}{2}$, a drawn-out S-shaped form somewhat similar to that of x itself, but not usually symmetrical.

From equation 63 it will be seen that f *may* have two changes of curvature as x varies from x_0 to L, namely, at the two roots of the quadratic expression in the value of d^2x/dt^2. The roots are

$$6x = (2\mathrm{L} + 1) \pm \sqrt{[(2\mathrm{L} + 1)^2 - 6\mathrm{L}]}; \qquad (65)$$

and are real and positive. The lesser root is always less than L, and is therefore always attained by x as it varies from x_0 to L—so that f always has at least one change of curvature. The greater root is also less than L when $\mathrm{L} > \frac{1}{2}$; is equal to L when $\mathrm{L} = \frac{1}{2}$; but is greater than L when $\mathrm{L} < \frac{1}{2}$—so that the second change of curvature occurs only when $L > \frac{1}{2}$. As x never quite reaches L, there is no second change of curvature when $L = \frac{1}{2}$. These results are therefore as were to be expected from the two forms of the curve of f already discussed.

As already seen, when $L > \frac{1}{2}$, the summit of the f-curve is $\frac{1}{4}c$ and is reached when $x = \frac{1}{2}$. The corresponding value of t is

$$t = \tau - 1/\mathrm{KL} \,.\, \log(2\mathrm{L} - 1), \tag{66}$$

the logarithm being negative since $2L - 1$, that is $1 - 2(N + r)/K$, is less than unity.

It has also been seen that when $L > \frac{1}{2}$, f falls after reaching its maximum, $\frac{1}{4}c$, to a constant value, $c\mathrm{L}(1 - \mathrm{L})$. The difference between the maximum and the ultimate values of f is therefore

$$\tfrac{1}{4}c\,(2\mathrm{L} - 1)^2,$$

and their ratio is

$$1/4\mathrm{L}\,(1 - \mathrm{L}) = \lambda \quad \text{(say)}. \tag{67}$$

As with independent happenings (*e.g.*, equation 22), we shall use the symbol l to denote the limit which f approaches when t is very large, and this will be examined further in Section VIII (iii) (Part II).

(vi) The curves of new cases in proportional happenings are especially important because, if the original assumptions upon which they are based are sufficient in themselves to explain time-to-time variations in the frequency of infectious diseases, these curves should agree, at least for a first approximation to the truth, with the curves of new cases actually observed in such diseases. Now, according to Brownlee (Section I), the curves of many epidemics, especially of the short and sharp zymotic diseases, tend to be *remarkably symmetrical bell-shaped curves*, roughly similar, in fact, to the f-curve developed from certain values of the constants.

We now proceed, therefore, to examine these constants with greater care. They are in detail,

$$\mathrm{K} = c - v + \mathrm{V}, \quad \mathrm{KL} = c - v + \mathrm{V} - \mathrm{N} - r,$$
$$\mathrm{L} = 1 - (\mathrm{N} + r)/(c - v + \mathrm{V}). \tag{68}$$

The particular elements have been already discussed in Sections III and V. The natural-variation element v governs the natural increase of the population apart from exceptional happenings; but V, N, r are connected with the special happening under consideration. If the period covered by such happening is short compared with the average life of the individuals concerned, v may be so small as to be negligible, and N also may be very small. In injurious happenings such as infectious diseases, $v - V$ should be positive [Section V (iii)], and V will be negative if $M > N$ and $E > I$, so that $v - V$ still remains positive even when $v = 0$; but, as a rule, except in very fatal diseases, $v - V$ is very small. The element r may be unity when reversion is immediate, as in slight accidents, but may be nearly zero when the immunity conferred by one attack of a disease lasts for a long time on the average, as in many zymotic diseases. Both N and r are only positive quantities, so that $v - V + N + r$ is usually a small or very small positive fraction in the case of infectious diseases (its actual magnitude depending, of course, on the magnitude of the time-unit taken as well as on the elements themselves). The element of special importance is c, that is the average number of individuals infected or reinfected in unit of time by each already infected individual. But while v, V, N, r may be taken as fixed for the kind of happening considered, c is generally unknown, and, indeed, one of the ultimate objects of such studies as these will be to ascertain its value. At present, however, we take it to be a constant.

We observe first from equation 53 that if KL is negative x must constantly diminish as t increases, which means that a change must have occurred in at least one of the original constants. Hence we assume at present that KL is positive, and reserve the study of the case when it is negative for Section IX (Hypometric Happenings: Part II).

Secondly, as KL is taken to be positive, c, whatever it is, must be greater than $v - V + N + r$—generally a small positive fraction; and therefore greater than each item individually.

Thirdly, as $c > v - V + N + r$ and N and r are positive, $c > v - V$; that is, K is always positive.

Fourthly, as $KL > 0$, $K > N + r$, and therefore L is always less than unity (as it must otherwise obviously be), and always positive. Clearly also, L increases with the increase of c and V, but diminishes with the increase of v, r, and N (the last of which exists in K but cancels out from KL).

We must now examine the conditions which hold respectively if f is to be of Type II or Type I, or is to be nearly or quite symmetrical.

It will be of Type II, that is, always increasing in a drawn-out S-shape [(v) above] if L is not greater than $\frac{1}{2}$; that is, if $c \ngtr v - V + 2(N + r)$. But as just shown, in the cases which we are now considering, c must always be greater than $v - V + N + r$. Hence in order that f shall be of Type II, c must lie between the limits $v - V + 2(N + r)$ and $v - V + (N + r)$. Now if $N + r$ is small, as is usually the case in short and sharp epidemics of zymotic diseases with long immunity, these limits will be narrow; in other words, out of the whole range of possible values of c, that is of infectivity, only a small sector, and that the lowest one, will make f of Type II.

On the other hand, f will be of Type I, that is, of a bell-shape, regular or irregular, if $L > \frac{1}{2}$—that is, if c is any number greater than $v - V + 2(N + r)$. In this case f rises to a maximum ordinate $\frac{1}{4}c$ and then falls to a constant level $cL(1 - L)$. Lastly, f becomes nearly symmetrical (see (v) above) when $L = 1$ nearly, that is, when $c = v - V + (1 + \gamma)\ (N + r)$, where γ is considerable; and also of course when N and r are very small. These are probably just the conditions which hold in many of the short and sharp epidemics of zymotic diseases, such as measles, scarlatina, and dengue; and they produce curves which differ from the perfectly symmetrical curve of $c/K \,.\, dx/dt$ only by a small increasing term which, however, never exceeds $c/K \,.\, (N + r)L$.

For an example, suppose the case of a human infectious disease in which $v - V = 0{\cdot}002$, $N = 0{\cdot}0004$, $r = 0{\cdot}02$, all assumed for a time-unit of one week. Then KL will not be positive unless $c > 0{\cdot}0224$; that is, the happening will have no effect unless, say 1000 affected individuals infect or re-infect 22·4 individuals per week on the average. If 1000 such persons can infect or reinfect 42·8 persons per week or less, the f-curve will be of Type II and will always ascend to the limit $\frac{1}{4}c$ (that is, 0·0107 if $c = 0·0428$), at which figure it will remain indefinitely. If, however, c is greater than this — if the 1000 affected persons can infect or re-infect, say, 100 persons per week — then the f-curve will rise to a maximum value of 0·025 of the total population, and afterwards decline to the constant value of 0·01647 of the total population, or 16·47 per thousand, at which it will remain (L being 0·792 in this case). If c is still greater, if each affected person can affect one other person per week, so that $c = 1$, then L = 0·9795, that is, nearly the whole population ultimately becomes affected. Here f, the proportion of new cases, reaches a maximum of one quarter the whole population and thereafter declines to the constant value of 0·01997, or nearly 2 per cent. of the population, at which it remains indefinitely. (For further examples see Section IX, Part II.)

(vii) Integral expressions for P and Z were obtained in equations 51 and 52 by integrating both sides of equation 44. These may be written

$$\mathrm{P} = \mathrm{P}_0\varepsilon^{vt}(x/x_0 \,.\, \varepsilon^{-\mathrm{KL}t})^{(v-\mathrm{V})/\mathrm{K}} = \mathrm{P}_0(x/x_0)^{(v-\mathrm{V})/\mathrm{K}}\varepsilon^{vt}\varepsilon^{-(v-\mathrm{V})\mathrm{L}t}, \qquad (69)$$

$$\mathrm{Z} = \mathrm{Z}_0\varepsilon^{vt}\varepsilon^{\mathrm{KL}t}(x/x_0 \,.\, \varepsilon^{-\mathrm{KL}t})^{c/\mathrm{K}} = \mathrm{Z}_0(x/x_0)^{c/\mathrm{K}}\varepsilon^{vt}\varepsilon^{-(v-\mathrm{V})\mathrm{L}t}. \qquad (70)$$

Two important special cases present themselves here. If $v - \mathrm{V} = 0$, the happening has no effect on the total number of the population, as in mild infectious diseases such as chicken pox or (?) dengue or party politics. This should be called the Equivariant Case of Proportional Happening. In it, $\mathrm{P} = \mathrm{P}_0\varepsilon^{vt}$ as in Section IV (ii), and of course $\mathrm{Z} = \mathrm{Z}_0\varepsilon^{vt}$. But if $v - \mathrm{V}$ is small without being zero, care must be taken with the functions if c, N, and r are small.

The other special case occurs when $v = 0$, and may be called the Case of Constant Natural Population. It is particularly important, since happenings which are considered during only a short period compared with the average length of life of the individuals concerned approximate to this type, because then v, the element of natural increase, is generally so small compared with the other constants that it may be neglected; in fact, the *a priori* curves of most epidemic diseases may be considered on this basis.

If $v = 0$ and V is negative (as in all injurious happenings) we see that, since

$$P = P_0\,(x/x_0)^{-V/(c+V)}\,\varepsilon^{VLt}, \tag{71}$$

the population diminishes indefinitely as time progresses; which is, of course, what was to be expected since by supposition there is no natural increase to compensate for the loss due to the happening.

If neither $v = 0$ nor $v - V = 0$ then, since $x = L$ ultimately, that is, a constant, the question whether the population will finally increase or diminish depends upon the sign of the expression

$$v - (v - V)\,L = V + [(v - V)\,(N + r)]/(c - v + V). \tag{72}$$

An important value of P is reached when the x-curve attains its centre of symmetry, that is when $x = \frac{1}{2}L$ and

$$t = \tau = 1/KL\,.\,\log\,[(L - x_0)/x_0],$$

(see (iii) above). Then

$$\begin{aligned} P &= P_0\,(L/2x_0)^{(v-V)/K}\,[(L - x_0)/x_0]^{[v-(v-V)L]/KL}, \qquad (73)\\ &= P_0\,(\tfrac{1}{2})^{(v-V)/K}\,(LP_0)^{v/KL}, \end{aligned}$$

if x_0 is very small, as when it equals $1/P_0$.

Another important value of P is reached when $Z = LP - 1$, that is when $t = 2\tau$ (equation 58). Then

$$P = P_0\,[(L - x_0)/x_0]^{[2v-(v-V)L]/KL}. \tag{74}$$

A third important value of P is reached (if $L > \frac{1}{2}$) when the f-curve attains its maximum; that is, when $x = \frac{1}{2}$ and $t = \tau'$ (equation 66). Then

$$P = P_0(1/2x_0)^{(v-V)/K}\,[1/x_0\,.\,(L - x_0)/(2L - 1)]^{[v-(v-V)L]/KL}. \tag{75}$$

These expressions are required to estimate the actual values of Z and F at these points: but, if L is not very small, τ will not be very large [see Section VIII (vii)], and therefore P can vary but little in that period from its original P_0, and, at least for rough calculations, can be taken as remaining constant.

(viii) It remains to consider the integrals of P, Z and F, which will be required for the analysis of several questions. We proceed as in the case of constant happenings [Section IV (v)] by integrating both sides of the fundamental differential equations (44 and 45). Hence

$$P - P_0 = r\int_0^t P\,dt - (r - V)\int_0^t Z\,dt. \tag{76}$$

$$Z - Z_0 = \int_0^t F\,dt + (V - N - r)\int_0^t Z\,dt. \tag{77}$$

To interpret the first of these equations, we note that the actual change in the total population effected during the period t, namely, $P - P_0$, equals what the change would have been by natural-variation only, namely $\int_0^t rP\,dt$, less the difference between the natural-variation and the happening-variation in the affected part of the population, namely $\int_0^t (r - V)Z\,dt$. The second equation, which is the same as equation 24, has been already interpreted in Section IV (v). It means that the total change in the number of the affected during the period t, namely, $Z - Z_0$, is equal to the sum of new cases, namely $\int_0^t F\,dt$, plus the affected immigrations, and less the affected emigrations, deaths, and recoveries, namely,

$$\int_0^t (I - E - M - r)\,Z\,dt.$$

Both these interpretations agree with what was to be expected.

The integrals of F and Z can be obtained at once in the important case when $c = 0$; and are

$$\int_0^t Z\,dt = (P - P_0)/V. \tag{78}$$

$$\int_0^t F\,dt = Z - Z_0 - (V - N - r)(P - P_0)/V. \tag{79}$$

If in these we have also $V = 0$, then $(P - P_0)/V$ becomes indeterminate since P remains constant: but the value of the fraction then is, from equation 71,

$$(P - P_0)/V = Lt - \frac{1}{c}\log_e(x/x_0). \tag{80}$$

The integrals of Z and F can also be easily obtained by direct integration of their values when $c = 0$.

We should note that the total births, deaths, immigrations, emigrations, and recoveries among the affected during the period t are given, as with constant happenings, by multiplying $\int_0^t Zdt$ by N, M. I, E, and r respectively.

The main result has already been summarised in the prefatory section. We can see from (vi) above that, for many values of the constants, the solution of the fundamental proposition stated in Section II yields curves generally similar to the curves frequently found in epidemics—that is, curves which are bell-shaped and nearly symmetrical, and tend to decline more slowly than they rose. This decline is in the case of the *a priori* curves due merely to the exhaustion of susceptible material; but the further studies contained in the second part must be considered before attempts are made to apply the results to actual observations.

REFERENCES

1. William Farr, “On the Cattle Plague,” ‘Daily News,’ 1866, February 17, and ‘Journ. of Social Science,’ London, 1866, March 20.
2. G. H. Evans, “Some Arithmetical Considerations on the Progress of Epidemics,” ‘Trans. Epidemiolog. Soc. of London,’ vol. 3, Part III, p. 551 (1873–75).
3. J. Brownlee, “Studies in Immunity: Theory of an Epidemic,” ‘Proc. Roy. Soc. Edin.,’ 1906, vol. 26, Part VI, p. 484.
4. J. Brownlee, “Certain Considerations on the Causation and Course of Epidemics,” ‘Proc. Roy. Soc. Med. London,’ 1909, June.
5. J. Brownlee, “The Mathematical Theory of Random Migration and Epidemic Distribution,” ‘Proc. Roy. Soc. Edin.,’ 1910, vol. 31, Part II, p. 261.
6. J. Brownlee, “Periodicity in Infectious Disease,” ‘Proc. Roy. Phil. Soc. Glasgow,’ 1914.
7. J. Brownlee, “Investigations into the Periodicity of Infectious Diseases by the Application of a Method hitherto only used in Physics,” ‘Public Health,’ London, 1915, vol. 28, No. 6.
8. J. Brownlee, “On the Curve of the Epidemic,” ‘Brit. Med. Journ.,’ 1915, May 8, p. 599.
9. Ronald Ross, ‘The Logical Basis of the Sanitary Policy of Mosquito Reduction,’ Congress of Arts and Sciences, St. Louis, U.S.A., vol. 6, p. 89 (1904); also ‘Brit. Med. Journ.,’ 1905, May 13.
10. Ronald Ross, ‘The Prevention of Malaria in Mauritius,’ Waterlow and Sons, London, 1908, pp. 29–40.
11. Ronald Ross, ‘The Prevention of Malaria,’ John Murray, London, 1910; also 2nd Edition, 1911, ‘A Special Addendum on the Theory of Happenings,’ pp. 651–686.
12. Ronald Ross, “Some Quantitative Studies in Epidemiology,” ‘Nature,’ London, 1911, October 5. Italian translation, ‘Malaria,’ Rome, 1911, November.
13. Ronald Ross, “Some *a priori* Pathometric Equations,” ‘Brit. Med. Journ.,’ 1915, March 27.
14. H. Waite, “Mosquitoes and Malaria,” ‘Biometrika,’ London, 1910, October, vol. 7, No. 4, p. 421.

PARTS II and III

PREFACE

Part I of this paper was published in the "Proceedings," 1916, A, vol. 92, having been read on November 11. 1915. In June 1915, the Royal Society was kind enough to give a Government Grant for providing me with assistance in order to complete the paper, and for carrying on further studies upon the subject; and Miss Hilda P. Hudson, M.A., Sc.D., was appointed for the work from May 1, 1916. The continuation of the paper has accordingly been written in conjunction with her; and I should like to take the opportunity to express my obligations to her for her valuable assistance, especially in regard to Part III.

The entire paper here presented is still limited to the theoretical side of the subject, as defined in the Introduction to Part I. Records of epidemics are now being examined in order to find how far the theoretical results which we have reached may be applied to them; but these studies must be reserved entirely for future discussion. [They were never completed.]

RONALD ROSS

SECTION VIII

(1) *Hypothetical Epidemics: Deduction of Constants.*—In any case of independent happening we can use equations 8 and 12 with V, v, N, n, all zero.

Then $$dx/dt = h\,(1 - x), \qquad \text{and, if } x_0 = 0,$$

$$\begin{aligned} ht &= -\log_\epsilon (1 - x), \\ &= x + \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 + \ldots \text{ if } x \text{ is small.} \\ &= 2\cdot30 \log_{10} \frac{1}{1 - x} \text{ for all values of } x. \end{aligned}$$

If the value of x is known for any value of t, say a year, this equation enables us to find h, and thence x for any other period, hour, day, week, etc.

Thus suppose that child-birth, death, or migration occurs to 20 *per mille* of a population of all ages during one year, to what proportion do they occur in one day? We have

$$\begin{aligned} 365\,h &= 0\cdot02 + 0\cdot0002 + 0\cdot0000027 + \ldots. \\ h &= 0\cdot000055. \end{aligned}$$

To find the proportion for one hour, we divide by 24: and to find it for 7 or 30 days, we multiply by these numbers respectively.

Or suppose that 85 per cent. of plague cases, taken at any stage of the illness, die within 30 days, and we require the daily death-rate of such mixed cases. Then we have

$$\begin{aligned} 30\,h &= -\,2\cdot30 \log_{10} (1 - 0\cdot85). \\ h &= 0\cdot063, \end{aligned}$$

giving a constant death-rate of 6·3 per cent. *per diem.* If the cases are all taken from the commencement of the illness, the curves are quite different, and show the highest death-rate some days after the onset—but this does not concern us at present.

If the annual birth-rate is 30 *per mille*, and the annual death-rate is 20 *per mille*, then the corresponding daily rates are

$$n = 0\cdot000083 \qquad \text{and} \qquad m = 0\cdot000055.$$

Hence if the immigration and emigration rates counterbalance each other we may take for the daily variation element

$$v = n - m = 0\cdot000028.$$

In this example the increase of the population at the end of the year would amount to 1 per cent.—that is, P would vary from P_0 at the beginning of the year to $1\cdot01\ P_0$ at the end of it. But if we use the value of v just given for equation 13, namely, $P = P_0\varepsilon^{vt}$ (putting $t = 365$ and $vt = 0\cdot0103$), we obtain $P = 1\cdot0103\ P_0$ at the end of the year—which is a little greater than the figure just given. The reason for this is, of course, that in equation 13 we assume that the propagation is continuous—that is, that during the year, the progeny of the original population will themselves have progeny. But this is not assumed in the value of P first obtained.

We may also calculate v from equation 13 if we know the values of P_0 and P at the beginning and end of t. Thus the population of England and Wales was 10,164,256 in 1811, and 36,070,492 in 1911. Then since $vt = \log_e P - \log_e P_0$ and $t = 365 \times 100$, the daily variation-element is $v = 0\cdot0000347$, the yearly variation-element is $0\cdot012666$, and the ten-yearly variation-element $0\cdot12666$. If a population doubles itself in a period t, we have $vt = \log_e 2 = 0\cdot693$, so that at the above rate the English population ought to double itself in 19,975 days, or nearly 55 years—which is about the case.

If the daily death-rate of plague is taken at the figure just calculated, of $0\cdot063$, and the plague birth-rate is a normal one of 30 *per mille* a year (so that $N = n$), then $V = N - M = -0\cdot063$; so that, if $v = 0\cdot000028$, $v - V = 0\cdot063 \ldots$ daily. Such figures give us some concrete ideas of what the constants are likely to be in cases of human epidemics. [Compare Sections III and V (iii).]

The reversion-element r is particularly elusive since we seldom have known figures to deal with regarding it. It is best to proceed as we did in connection with equation 20, and to assume that 90 per cent. of the affected individuals revert in the time t, where $t = 1, 2, 3, \ldots, 10$, or more years.

The reader is reminded that KL is the coefficient of l in x, the proportion of affected individuals (equation 50), and that the x-curve reaches its centre of symmetry when $x = \frac{1}{2}$ L, and its maximum value L when t is very large [Section VII (iii)]. The period at which x reaches its centre of symmetry and dx/dt reaches its maximum is $t = \tau = \frac{1}{\mathrm{KL}} \log_e \frac{\mathrm{L} - x_0}{x_0}$. If $x_0 = 1/\mathrm{P}_0$ and LP_0 is large compared with unity, then, as suggested in connection with equation 58, we may now write

$$\gamma \mathrm{R} \tau = \log_e \mathrm{P}_0 + \log_e \gamma - \log_e (\gamma + 1),$$

and at twice this period the maximum number of affected individuals less one is reached.

Similarly, when t is very large, the ultimate value of the proportion of new cases is l. But if $\mathrm{L} > \frac{1}{2}$, f rises to a previous maximum value $\frac{1}{4}c$ when $x = \frac{1}{2}$, and thereafter falls to its ultimate value l [Section VII (v)]. The ratio between the maximum and the ultimate value is λ (equation 67). The period at which the maximum value is reached is

$$t = \tau' = \frac{1}{\mathrm{KL}} \log_e \frac{1}{x_0} \frac{\mathrm{L} - x_0}{2\mathrm{L} - 1}$$

(equation 66); and when $x_0 = 1/\mathrm{P}_0$ and LP_0 is large, this may be written

$$\gamma \mathrm{R} \tau' = \log_e \mathrm{P}_0 + \log_e \gamma - \log_e (\gamma - 1).$$

But this maximum of f (previous to its ultimate value) occurs only if $\gamma > 1$.

It is evident that when c is expressed as a function of γ in this manner, K, KL, L, τ, τ', and λ are independent of v and V, and that l contains them only as parameters of c.

On the other hand, τ and τ' are functions of the original population P_0 and increase as it increases—because, obviously, x and f require a longer period to reach their corresponding values if the population is large. It is convenient therefore to adopt a definite figure for P_0 in the following Table; and we put $\mathrm{P}_0 = 10{,}000$, which is, say, the population of an ordinary town. We may also assume that unity is negligible compared with LP_0 unless $\gamma < 0{\cdot}1$.

Then for a daily element $365\, r = \frac{1}{t} 2 \cdot 30$. Hence if 90 per cent. of the cases revert in one year, the daily reversion-element is $r = 0 \cdot 00631$. Thus the reversion-element is likely to be large compared with v or N and correspondingly important, and will be by no means negligible even when the considered disease is one in which the acquired immunity is supposed to be very lasting [compare Section III (iii)]. For a case reverts not only when it becomes again capable of showing a recognisable infection, but when it is again able to harbour the infective agents sufficiently to afford them a *nidus* whence they may infect others; and it is quite possible and indeed probable that this may often occur much earlier than we imagine. Thus persons vaccinated against smallpox may acquire the disease in a very mild and modified form after only a few years, and may then infect non-protected persons in full force; and there is no proof that a mild first attack of measles or other diseases may not give quite as short-lived an immunity—that is, that those who have had the disease in infancy may not comparatively early acquire it again in an indistinguishable form and then spread it. In such cases, r may be considerably larger than we might otherwise expect.

Regarding immigration and emigration, it will suffice to note that in 1911 there were 350,429 immigrants into the United Kingdom and 454,527 emigrants from it.

Further remarks on the constants will be found in Section X.

(ii) We may now proceed to consider numerical examples of proportional happenings. For this purpose it is convenient to adopt the suggestion made towards the end of Section VII (vi), and to take c as a function of an independent parameter γ which is such that when γ is positive KL is positive and x always increases with the time, and when γ is negative KL is negative and x always diminishes with the time (Section IX). We shall also write D for $v - \mathrm{V}$ and R for $\mathrm{N} + r$. The principal constants of x and f may then be written

$$c = \mathrm{D} + (\gamma + 1)\,\mathrm{R}, \qquad \mathrm{K} = (\gamma + 1)\,\mathrm{R}, \qquad \mathrm{KL} = \gamma \mathrm{R},$$

$$\mathrm{L} = \frac{\gamma}{\gamma + 1}, \qquad l = c\,\frac{\gamma}{(\gamma + 1)^2}. \qquad \lambda = \frac{(\gamma + 1)^2}{4\gamma}.$$

If P_0 is not 10,000, but some multiple or fraction of it, say 10,000 p, all we have to do to find τ and τ' is to add to or subtract from the figures given in the corresponding columns of the table the quantity $\log_\epsilon p/\gamma R$. Thus, if $P_0 = 1000$, we subtract $2 \cdot 30/\gamma R$. As γ increases above 100, τ and τ' approximate more and more closely to the value $\log_\epsilon P_0/\gamma R$, τ being always less than this and τ' greater than it.

The actual curves depend largely on the values of R. If we suppose the case of a human disease, of which the affected birth-rate N equals the natural birth-rate n, which is taken at $0 \cdot 000083$ daily, corresponding, according to the previous sub-section, to an annual birth-rate of 30 *per mille*, then R equals the former figure if r is so small as to be negligible. But if r is a daily rate corresponding to an annual reversion of 90 per cent., it will amount to as much as $0 \cdot 0063$, so that $R = 0 \cdot 0064$. We will suppose that these two values of R are a minimum and a maximum.

Now if we suppose also that $c = 1$, that is, that each affected individual daily infects or reinfects one other individual, then we have $\gamma = (1 - D)/R - 1$. That is, if $D = 0$, $\gamma = 12{,}000$ if R is a minimum, and $\gamma = 155 \cdot 5$ (or roughly 150) if R is a maximum. Thus, if the infection rate is near unity, the value of γ will reach from three to five figures, and the epidemic will resemble the last examples given in the Table.

In Example 1, if the population remains the same during the whole of this long period (which is, of course, never likely to occur), only about ten living persons out of the 10,000 will be found unaffected at the end of it; and the ultimate and maximum daily number of new cases lP_0 reaches only $0 \cdot 00083$, that is, only one new case in $3 \cdot 3$ years.

If, however, R has its maximum value, there is one new case in about sixteen days.

In Example 4, half the population is ultimately affected, but still f has no maximum before its ultimate value. When, however, in Example 5, γ increases above unity, this maximum begins to appear, and the ratio λ between the maximum and ultimate values of f constantly increases with γ. At the same time, τ and τ' diminish and approach each other.

Table of Hypothetical Epidemics.

Ex. No.	γ	L	$\frac{l}{c}$	λ	Reversion rate 90 per cent. per annum R = 0·0064				Reversion rate negligible R = 0·00083			
					c	τ	τ'	lP_0	c	τ	τ'	lP_0
1	0·001	$\frac{1}{1001}$	0·000998	—	0·0064	years 940	days —	0·064	0·00008	years 72,100	years —	0·00083
2	0·01	$\frac{1}{101}$	0·00980	—	0·0065	196	—	0·64	0·00008	15,100	—	0·0082
3	0·1	$\frac{1}{11}$	0·0826	—	0·0070	29	—	5·8	0·00009	2,200	—	0·075
4	1	$\frac{1}{2}$	$\frac{1}{4}$	1	0·0128	3·7	—	32	0·00017	280	—	0·41
5	2	$\frac{2}{3}$	$\frac{2}{9}$	$\frac{9}{8}$	0·0192	—	774	43	0·00025	—	163	0·55
6	3	$\frac{3}{4}$	$\frac{3}{16}$	$\frac{4}{3}$	0·0256	—	501	48	0·00033	—	105	0·62
7	4	$\frac{4}{5}$	$\frac{4}{25}$	$\frac{25}{16}$	0·0320	—	371	51	0·00042	—	78	0·66
8	5	$\frac{5}{6}$	$\frac{5}{36}$	$\frac{9}{5}$	0·0384	—	295	53	0·00050	—	62	0·69
9	6	$\frac{6}{7}$	$\frac{6}{49}$	$\frac{49}{24}$	0·0448	—	245	55	0·00058	—	51	0·71
10	7	$\frac{7}{8}$	$\frac{7}{64}$	$\frac{16}{7}$	0·0512	—	209	56	0·00067	—	44	0·73
11	8	$\frac{8}{9}$	$\frac{8}{81}$	$\frac{81}{32}$	0·0576	—	183	57	0·00075	—	38	0·74
12	9	$\frac{9}{10}$	$\frac{9}{100}$	$\frac{25}{9}$	0·0640	—	162	58	0·00083	—	34	0·75
13	10	$\frac{10}{11}$	$\frac{10}{121}$	3·03	0·0704	—	146	58	0·00092	—	31	0·76
14	100	$\frac{100}{101}$	$\frac{100}{101^2}$	5·5	0·646	—	14·4	63	0·00843	—	3	0·82
15	1,000	$\frac{1000}{1001}$	$\frac{1000}{1001^2}$	251	6·41	—	1·44	64	0·0835		days 110	0·83
16	10,000	$\frac{10000}{10001}$	$\frac{10000}{10001^2}$	2500	64·1	—	0·14	64	0·835		11	0·83

If c is nearly unity, $\tau' = 2 \cdot 30 \log_{10} P_0 = 9 \cdot 2$ days, and f reaches its maximum in a short time, whatever the value of R. More generally, if $D = 0$, $c = (\gamma + 1)R$, and if γ is large (say over 50), $\gamma + 1 = \gamma$ roughly, and $\tau = 1/c \,.\, \log_\epsilon P_0$. In other words, if γ is considerable, not only does $L = 1$ nearly, but the maximum ($\frac{1}{4}c$) of f and the mode, or time (τ') when that maximum is attained, depend roughly on c only, the population being fixed.

If $L > \frac{1}{2}$ the f-curve slopes downward from its summit more slowly than it rose towards that summit. Put $x = \frac{1}{2} + \xi$, then $f = c\,(\frac{1}{4} - \xi^2)$, and has the same value for equal and opposite values of ξ.

But

$$\frac{df}{dt} = 2c\,|\xi|\,\frac{d\xi}{dt}; \qquad \frac{d\xi}{dt} = \frac{dx}{dt} = \frac{K}{c} f - K\,(1 - L)\,(1 + \xi),$$

and is less when ξ is positive than when ξ has the equal and opposite value. Hence, at two points of the f-curve whose ordinates are equal, the descending side is less steep than the ascending side.

Since $f = \frac{c}{K}\frac{dx}{dt} + c(1 - L)\,x$, and the first term on the right gives a perfectly symmetrical bell-shaped curve, we may call $c\,(1 - L)\,x$ the *excess* of the f-curve. When $D = 0$ and γ is considerable, the excess varies roughly from $\frac{1}{2}R$ at the centre of symmetry to R, when t is very large—that is, roughly, from $\frac{1}{2}l$ to l. Both of these become small when r, the reversion-rate, is small.

The important result, therefore, follows that, if also c is near unity, and therefore γ is large—that is, if the infection-rate is high while the reversion-rate is low—then the f-curve becomes a nearly symmetrical bell-shaped curve. This, according to Dr. Brownlee (Section I), is just the kind of curve presented by epidemics of certain zymotic diseases. *The* prima facie *inference is therefore that such epidemics may be wholly or chiefly mere cases of proportional happening, as defined in Section VII* (i). It is even the case that, in such diseases, the fall of the curve is often more slow than its rise, as happens with the f-curve.

When γ is large and $L = 1$ nearly, the two changes of curvature of the curve occur, according to equation 65, when $6x = 3 \mp \sqrt{3}$, that is, when (equation 53)

$$\begin{aligned} \gamma Rt &= \log_e P_0 + \log_e (1/x - 1). \\ &= \log_e P_0 - 1\cdot32 \quad \text{and} \quad \log_e P_0 + 1\cdot03. \end{aligned}$$

the latter being nearer to the summit than the former.

(iii) Several methods may be employed for calculating the curves of x, f, Z and F in detail. The simplest is to divide the period during which x varies from x_0 to $\frac{1}{2}$L into a number (say 10) of equal parts and then to calculate the ordinate of the curve at each section. In equation 58 we used τ to express this interval when $x_0 = 1/P_0$, but it may be employed more generally to express the abscissa of the centre of symmetry, when x_0 has any value, so that $KL\tau = \log_e (L/x_0 - 1)$. We now put $t = \tau T/10$, give to T the successive values 1, 2, 3, ..., 10, and calculate the corresponding values of x, namely, x_1, x_2, x_3, ..., and so on. Thus we have from equations 50

$$x_T = L\{\sqrt[10]{(L/x_0 - 1)^{10-T}}\}^{-1}.$$

The root can be easily evaluated by means of logarithms, and the value of x then obtained from a table of reciprocals (*e.g.*, as in Barlow's Tables). Owing to the symmetry of the x-curve, its values when T lies between 10 and 20 can be easily found—for example $x_{13} = L - x_7$. When we have ascertained the ordinates of x, we can quickly calculate those of f from equation 61, and, unless the case-mortality is high, we can generally assume for short and sharp epidemics that the original population remains constant and that $Z = xP_0$ and $F = fP_0$. It will be observed that, by this method, when τ has been first calculated we need subsequently deal only with the values of L and of $L/x_0 - 1$. Thus suppose that L = 1 nearly, that P = 10,000, and that $x_0 = 1/P_0$, then we have approximately,

T = 0	1	2	3	4	5	6	7	8	9	10
Z = 1	2·5	6·3	15·8	39·7	99	245	593	1368	2847	5000
F/c = 1	2·5	6·3	15·8	39·6	98	239	558	1181	2037	2500

T =	11	12	13	14	15	16	17	18	19	20
Z =	7153	8632	9407	9755	9901	9960	9984	9994	9997	9999
F/c =	2037	1181	558	239	98	39·6	15·8	6·3	2·5	1

This is, of course, an ultimate case where f is quite symmetrical—so that, for example $f_7 = f_{13} = cx_7x_{13}$.

Another useful method for calculating x and f is to transfer the origin to the centre of symmetry (equation 56). But if the ordinates are required for successive natural units of time, such as days or weeks, in order to compare them with statistics of epidemics in which such units have been used, then we must of course give to t the successive values 1, 2, 3, ..., and calculate the ordinates directly from equations 50. Here the symmetry of x will not be so useful, because the centre of symmetry is not likely to coincide with any one of the periods of time for which the ordinates have been found. It will, therefore, be generally easier to plot the curve by the method first indicated, to measure off the natural units of time along the time-axis, and then to estimate the corresponding ordinates geometrically.

SECTION IX

(i) *Hypometric Happening*: KL *Negative.*—We have hitherto always supposed that the happening-element is constant during the whole of the considered period, but it is now necessary to examine the changes which will be produced in the functions by certain changes in this element. We here study only the cases of dependent happenings already discussed, and these only for hypometric happening. It is no longer very useful to take c as a function of γ.

In Section VII (vi), while considering the constants K, L, c, we saw from equations 50 that if KL is positive, x always increases and is therefore always greater than x_0, its value when $t = 0$. This implies that $c > D + R$. When this inequality becomes an equality, KL vanishes and $dx_0/dt = 0$, that is, x remains near its original value when t is not very large, so that the happening may be said to be *isometric* [see (v) below]. When, however, $c < D + R$, KL becomes negative; and it is seen from equations 50 that, as ε^{-KLt} then always increases with the time, x always diminishes, and must therefore be always less than x_0.

We call such cases *hypometric happening*. That is, we now suppose that, after having been greater than $D + R$, c becomes less than it, or even zero. What, then, will be the nature of the curves?

The solutions of equations 44, 45, 47, 51, 52 remain valid for all real values of c, and we have therefore only to interpret results already known.

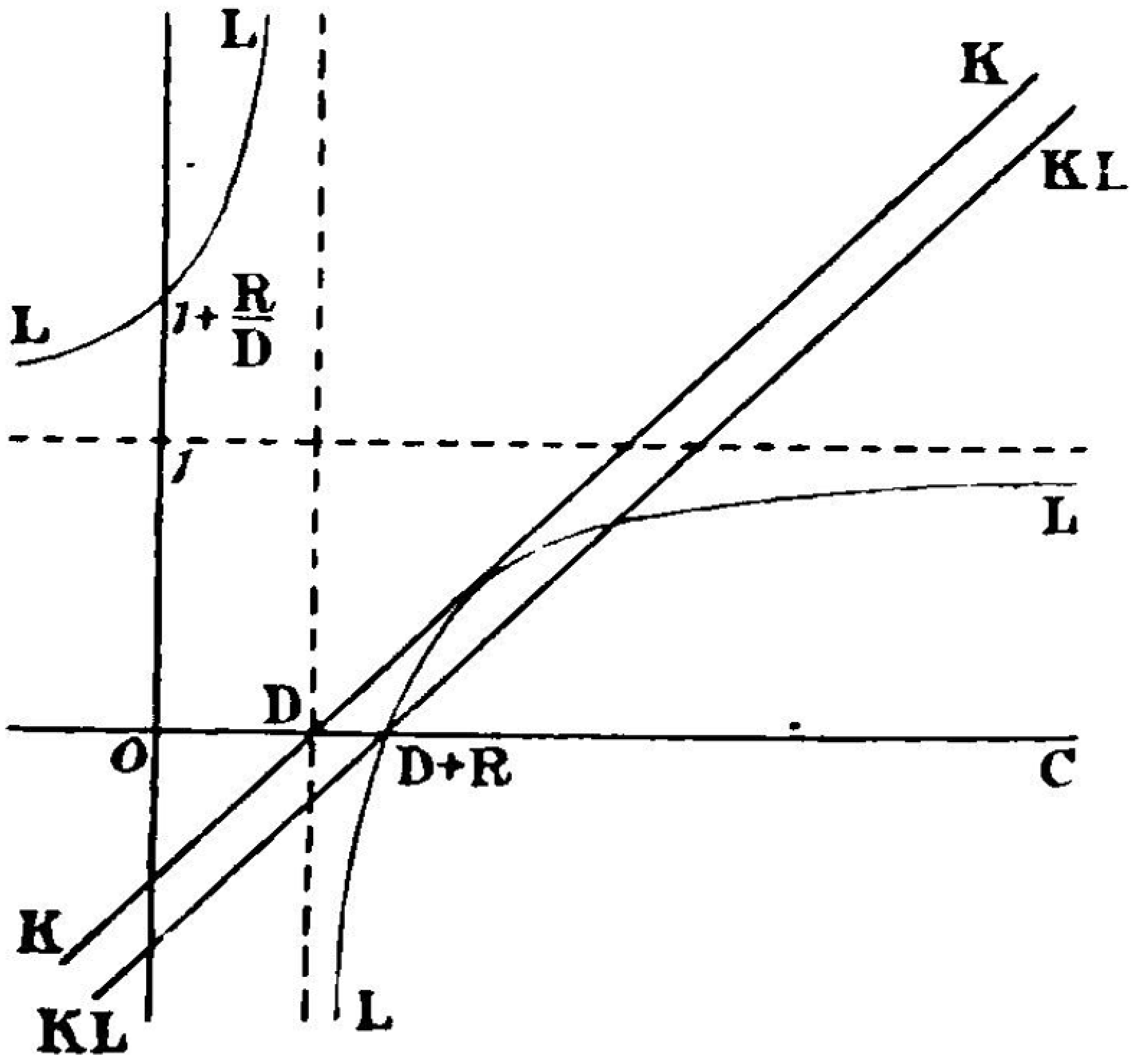

FIG. 1

It is first necessary to examine K, KL, and L more particularly as functions of c, but we shall do so only in the case when D is positive (injurious happenings) or zero. Then D, R are positive constants. We have

$$K = c - D, \qquad KL = c - D - R,$$

represented by parallel straight lines meeting the axis of c at an angle of 45° on the positive side of the origin.

$$L = 1 - R/(c - D), \quad \text{or} \quad (L - 1)(c - D) = - R,$$

and the graph of L is a rectangular hyperbola, with asymptotes $L = 1$, $c = D$, dotted lines in the figure, parallel to the axes of coordinates.

The last portion of this L curve, for $c > D + R$, gives cases of hypermetric happening already studied; but we have to consider the range $0 \leqq c < D + R$; since L becomes infinite when $c = D$, we have to consider the two portions: $0 \leqq c < D$, for which L is positive and greater than 1: and $D < c < D + R$, for which L is negative.

(ii) First let $c = 0$: that is, suppose that the happening suddenly ceases altogether, while all the other elements remain the same. It is advisable to use marked letters for the quantities after this event. We take it then that when the original happening c ceases, $t' = 0$, and that the values which x, Z, P, had reached up to this time are now expressed by x_0', Z_0', P_0', while x', Z', P' denote these functions as t' increases. Similarly, K′ and L′ are now the values of K and L without the element c. Then, since $K' = -D$, equations 51 and 52 become

$$P' = P_0' \varepsilon^{rt'} \left\{ 1 - \frac{x_0'}{L'} \left(1 - \varepsilon^{-(D+R)t'} \right) \right\}. \tag{81}$$

$$Z' = Z_0' \varepsilon^{(V-R)t'}. \tag{82}$$

As $V - R = -M + I - E - r$, Z' decreases from its original value Z_0' or Z, and ultimately approaches zero—unless the immigration is greater than the sum of the other elements (as may happen in a war area)—and its diminution is more rapid when M, E, r are large. With regard to equation 81, we observe that $P_0' \varepsilon^{rt'}$ is the natural increase of the population, living when the happening ceased, after that event: and the equation shows that P′ equals this quantity less a function of t' which is zero when $t' = 0$, which increases with t', and which finally slowly approaches the limit $x_0'/L' \,.\, P_0' \varepsilon^{rt'}$ when t' is large. That is to say, the population which remained when the happening ceased continues to suffer from its effects for some time afterwards, and then finally reaches the value

$$P' = P_0' \varepsilon^{rt'} (1 - x_0'/L').$$

Here x_0'/L' is always less than unity, for x_0' is the value of x at the moment when the happening ceased, and this could not have been greater than L, which is less than unity [Section VII (vi)], while, as shown in the previous sub-section, L′ is greater than unity when $c = 0$. Hence P′ is always less than $P_0' \varepsilon^{rt'}$, and the difference gives the loss of population due to the cases which continue after the happening ceased.

The fraction L'/x_0' occurs in the denominator of x' in equation 50. As it is greater than unity, the coefficient of $\varepsilon^{-K'L't'}$ is positive, so that x' always diminishes as t' increases—as could have been already inferred from equations 81 and 82.

(iii) The next case is when c' is some quantity between zero and D (which is generally very small). For the study of this case, the first of equations 50 may be written

$$x' = x_0' \left\{\frac{x_0'}{L'} + \left(1 - \frac{x_0'}{L'}\right)\varepsilon^{-K'L't'}\right\}^{-1}. \tag{83}$$

which diminishes indefinitely as t' increases. If $c' = D$ and L' is consequently infinite, this becomes $x_0\varepsilon^{K'L't'}$, which vanishes.

The values of P' and Z' are easily obtained from equations 51 and 52, if we remember that the indices D/K' and c'/K' are now negative. When $c' = D$, the expressions for P' and Z' become indeterminate, but can be easily evaluated, and we have

$$P' = P_0'\varepsilon^{vt'-q}, \quad \text{where} \quad q = x_0'D\,(1 - \varepsilon^{-Rt'})/R.$$

(iv) When c' lies between D and D + R, K' is positive and L' is negative, and varies from $-\infty$ to zero. Then it will be seen from equation 83 that x' is still positive, and, as in the previous cases, diminishes indefinitely as t' increases. For it may be written

$$x' = x_0' \left\{\varepsilon^{-K'L't'} - \frac{x_0'}{L'}\left(\varepsilon^{-K'L'T'} - 1\right)\right\}^{-1}, \tag{84}$$

where $K'L'$ is negative. For P' we have

$$P' = P_0'\varepsilon^{vt'}\left\{1 - \frac{x_0'}{L'}\left(1 - \varepsilon^{K'L't'}\right)\right\}^{-D/K'}, \tag{85}$$

which is also always positive since L' is negative, and increases without limit when t' is large. P' is less than $P_0'\varepsilon^{vt'}$, because the effects of the original happening continue to be felt by the population after the happening itself has been reduced from c to c'.

(v) When $c' = \mathrm{D} + \mathrm{R}$ exactly, L' and $\mathrm{KL}' = 0$, and, when t' is small, $x = x_0'$ nearly, and the happening may be said to be *isometric*. The values of x' and P' become indeterminate, but can be ascertained by finding the limits when $\mathrm{L}' = 0$ of the expressions in equations 84 and 85, so that

$$x' = \frac{x_0'}{1 + x_0'\mathrm{K}'t'},$$

$$\mathrm{P}' = \mathrm{P}_0'\varepsilon^{rt'}(1 + x_0'\mathrm{K}'t')^{-\mathrm{D}/\mathrm{K}'}.$$

Or the same equations can be obtained by integrating equation 47 after putting $\mathrm{L} = 0$ in it, and then integrating both sides of equation 44. In this case also, then, x' tends to decrease as t' increases, and finally approaches zero, when t' is very large.

(vi) If $\mathrm{K}'\mathrm{L}'$ is positive, however small it may be, x' always increases, as already seen in Section VII (iii).

It must be remembered that, though x', the proportion of affected individuals, diminishes, this does not mean that new cases cease to occur. On the contrary, unless c is absolute zero, f' is the same function of x' as before. And rx' expresses, as before, the reversions of the old cases.

When $\mathrm{D} = 0$, the phase described in (iii) above is suppressed, and we begin, when $c' = 0$, at the point when $\mathrm{K}' = 0$, $\mathrm{K}'\mathrm{L}' = -\mathrm{R}$, and $\mathrm{L}' = -\infty$, and the functions will be easily understood.

SECTION X

(i) *Parameter Analysis.*—We have now considered x, f, and P as functions of the time, but it remains to examine how they vary if we give different values to the parameters h, r, v, V, n, m, i, e, N, M, I, E. The symbol l has been used to denote the limit of f when t is very large. This is an important quantity in disease happenings because it gives the number of new cases which continue to occur when x has reached its limit, that is, after epidemic manifestations have ceased. It is, by hypothesis, these final new cases which keep the infection alive permanently in the population; we may call l the *endemic ratio.* First consider l as a function of r

$$\frac{\partial l}{\partial r} = \frac{c}{K}(2L - 1) = \frac{c}{K^2}(c - D - 2R).$$

Thus l is positive when $r = 0$ (if $KL > 0$) and increases with r until this reaches a value which makes $L = \frac{1}{2}$ and $l = \frac{1}{4}c$ [Section VII (vi)]. After this l diminishes as r increases. But if $c - D > 2R$ then r never reaches this value and l always increases. Thus l considered as a function of r has forms similar to those of f considered as a function of t.

The behaviour of l as a function of c is of the same nature. A similar procedure applies to f and P.

The effect upon the total population P due to variations in the parameters is important. The fundamental equation is 44:

$$dP/dt = vP - DxP.$$

As P and x are always positive, the sign of dP/dt depends on v and D. If $D = 0$, the happening is equivariant and the natural change of population due to v is not affected. If D is negative, the happening is beneficial and the natural change of population is augmented by it. If D is positive, the happening is injurious and the natural change is reduced. If $v - Dx$ is negative, especially when $x = L$, the total population will diminish indefinitely.

It is found that $\partial P/\partial r$ is always positive; that is, in infectious diseases, a quicker loss of affectedness, including loss of immunity, is beneficial. This may seem surprising until we reflect that, as $v > V$, the longer the affectedness continues in the individual the greater will be the loss of life due to it, since we generally suppose in such cases that $M > m$ and that these elements act during, not one, but every time-unit lived by the affected part of the population. Of course, the greater r is the more quickly on the average will affectedness, including immunity, be lost.

Next, $\partial P/\partial c$ is zero or negative, and the higher the infection rate the more injurious is the happening.

$\partial P/\partial V$ is zero when $t = 0$. It generally increases at first with t and afterwards decreases: it is negative when $x = L$ provided $l - c(1 - L)/K$ is negative. In other words, in these cases an increase of V is beneficial at the beginning of the outbreak, but injurious later on. Now $V = N - M + I - E$ and diminishes if the case-mortality M increases; in fact $\partial P/\partial M = -\partial P/\partial V$. Thus an increase in the case-mortality is injurious at first, but may be ultimately *beneficial*, because an affected individual who dies ceases to be infective.

PART III

SECTION XI

(i) *Variable Happening.* In Section IX. we commenced by touching upon the necessity of studying the effect of changes in the happening-element, but there dealt only with the case of hypometric happening. We now proceed to examine other changes in this element.

It is unlikely that in any infectious disease the infectivity-element c will always remain constant. It may remain practically such for some length of time: and the hypothetical results of this condition have therefore been analysed above in detail; but nature abhors a straight line, and we may infer from general experience that changes are sure to occur from time to time. As suggested in Section I, changes in infectivity may be due—

(*a*) to the action, so to speak, of the infecting organisms themselves —which is difficult to believe in because it involves the conception that organisms living in large numbers of different hosts should, as it were, make a kind of concerted and simultaneous effort; and

(*b*) to changes in season, climate, or other environment which may (1) increase or diminish the infective strength of the infecting organisms; or (2) facilitate or hinder their transference from one host to another.

At present we can proceed only on the supposition that the change has occurred, without attempting to enquire how or why it has occurred; and only some brief notes on the subject can be given at present.

(ii) *Discontinuous Change.*—We have hitherto supposed that when $t = 0$ the number of affected individuals is Z_0, which may be small or large. If the happening is hypermetric ($KL > 0$) and Z_0 is less than LP, the number of affected individuals, Z, will increase towards the latter figure; but if the happening is isometric ($KL = 0$), or is hypometric ($KL < 0$). Z will diminish to the limits discussed in Section IX. What will happen if the infectivity is suddenly changed during one unit of time from c to c', and thereafter remains at the latter figure?

If $c' > c$, the x-curve suddenly becomes steeper: x itself is continuous, but its gradient is discontinuous. The f-curve is discontinuous, and has a sudden rise at the moment when c changes: thereafter both curves proceed according to the equations with c' substituted for c. If at the moment of discontinuity more than half the population is affected, so that f is descending, there is an epidemic manifestation, a sudden rise followed by a decline. Both L and l are increased.

These are probably just the conditions which apply to many zymotic diseases, in which most of the population appears to remain affected, that is immune, for years or permanently. In these cases, at the moment of discontinuity x is nearly L, and L is nearly unity, and the augmented infectivity c' acts merely upon the small residue of non-affected individuals, let us say, newly-born infants, or children, or others who have hitherto escaped infection by accident or by residence in remote villages. It is well known that at the beginning of the war recruits from many highland villages were attacked by very virulent measles on joining camps in the lowlands. On the other hand, the same sort of epidemic may occur in numbers of villages and small towns in which the considered infection has died out from want of material, and so allowed of a gradual accumulation of fresh material, which is then suddenly fired (so to speak) by the entry of one or two infective cases from outside.

It is extremely improbable, of course, that the infectivity actually makes a single *per saltum* change; but we may still apply the equations already obtained if we give to c a series of progressive discontinuous augmentations for successive units of time, followed by similar discontinuous diminutions until c returns to its original value. In this case the f-curve will again show epidemic manifestations (if $x > \frac{1}{2}$), and will then decline again towards a lower value. This is one way of approximating to a continuous change in c.

(iii) *Continuous Variation.*—We may, however, obtain new equations on the hypothesis that the infectivity c is a continuous function of the time. Let us suppose that $c = \phi'(t)$, where $\phi'(t) = d\phi(t)/dt$. Then in the fundamental equation 47, namely, $dx/dt = \mathrm{K}x(\mathrm{L} - x)$, we have

$$\mathrm{K} = \phi'(t) - \mathrm{D} \quad \text{and} \quad \mathrm{KL} = \phi'(t) - \mathrm{D} - \mathrm{R}.$$

Put $x = 1/y$, so that $dy/dx = -\mathrm{KL}y + \mathrm{K}$; then

$$y \exp \int \mathrm{KL}\,dt = \int \Big(\mathrm{K} \exp \int \mathrm{KL}\,dt\Big)\,dt + \text{constant}$$

by the well-known solution. As however, $\mathrm{K} = \mathrm{KL} + \mathrm{R}$,

$$\int \Big(\mathrm{K} \exp \int \mathrm{KL}\,dt\Big)\,dt = \exp \int \mathrm{KL}\,dt + \mathrm{R} \int \Big(\exp \int \mathrm{KL}\,dt\Big)\,dt\,;$$

hence

$$y = 1 + \exp\Big(-\int_0^t \mathrm{KL}\,dt\Big)\Big\{\mathrm{R}\int_0^t \Big(\exp \int_0^t \mathrm{KL}\,dt\Big)\,dt + y_0 - 1\Big\}.$$

If R is negligible, as may often happen in short and sharp epidemics with long immunity, we may write, for many diseases at least,

$$\int \mathrm{KL}\,dt = \phi(t) - \mathrm{D}t,$$

$$x = \frac{1}{1 + \text{const.} \times \varepsilon^{-\phi(t) - \mathrm{D}t}}.$$

We have at present no knowledge of the nature of the function $c = \phi'(t)$ or even if there is such a function. But if there is, we may suppose that, for positive values of t, $\phi'(t)$ is small when t is small, increases at first as t increases, reaches a maximum for a certain value of t, and then declines to a small value again when t is larger—as is observed in many phenomena. In other words we should expect a symmetrical or non-symmetrical bell-shaped curve such as those with which we are familiar in statistical work.

(iv) *Graphical Treatment of x.*—We may also suppose that c is given in terms of the time t, not by an analytic function, but by a curve whose equation is unknown. We can then determine the nature of x and f by graphical methods.

Since $dx/dt = \mathrm{K}x(\mathrm{L} - x)$, there could only be a turning value of x when $\mathrm{K} = 0$ or $x = 0$ or $x = \mathrm{L}$. But when $\mathrm{K} = 0$, then $\mathrm{KL} = -\mathrm{R}$ and dx/dt does not vanish; when $x = 0$ the happening ceases altogether; the only turning value of x is when $x = \mathrm{L}$.

Now since by definition x lies between 0 and 1, we can only have $x = L$ if L also lies between 0 and 1; the diagram of IX (i) shows that this only occurs when $c > D + R$. Hence unless the given c-curve rises above the level D + R, the corresponding x-curve has no fluctuations, and either rises steadily to 1 as its limit, or sinks to 0. But wherever the c-curve has a portion rising above D + R, we can plot the values of $L = 1 - R/(c - D)$ corresponding to this portion, which form a cap, part of the L-curve, lying entirely between the levels 0 and 1. If the c-curve fluctuates, there are a series of these L-caps, which are the only parts of the L-curve that the x-curve can possibly cross.

Now, whenever the x-curve meets an L-cap, $x = L$, $dx/dt = 0$, and x has a turning value. If the x-curve descends to meet the L-cap from outside, it begins to ascend as soon as it is inside; and goes on ascending till it meets the opposite side of the same cap, when it passes outside and immediately begins to descend, until it meets

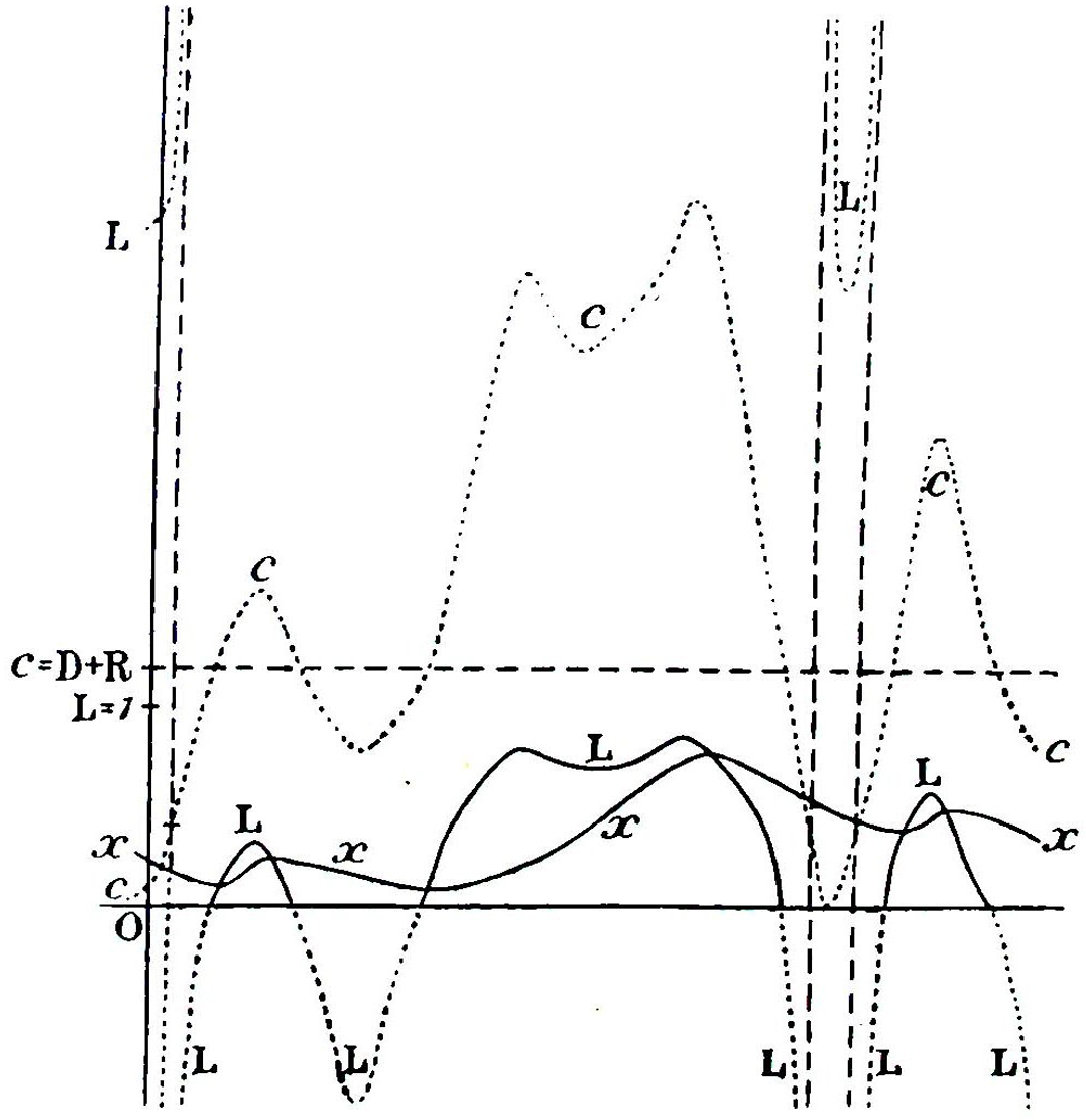

FIG. 1.

another L-cap. The x-curve descends steadily except where it is inside a cap; and as soon as the L-caps are plotted, the general character of the curve is fairly apparent. If c has a definite limit c_∞ when $t \to \infty$ and this limit is less than $D + R$, then the last L-cap is closed, x is ultimately descending and tends to 0. But if $c_\infty > D + R$, then L has a limit L_∞ and x tends to the same limit as L, and may approach it either from above or from below.

Since $L = 1 - R/(c - D)$, $\partial L/\partial c = R/(c - D)^2$, and is positive, L rises and falls with c and the top of an L-cap corresponds exactly to a maximum of c. But a maximum of x occurs where it crosses a descending branch of L: hence the maxima of x occur somewhat later than the corresponding maxima of c.

Suppose, for example, that c is periodic, rising and falling regularly every year, and crossing the level $D + R$ twice a year. Then L has a series of caps with equal equidistant maxima of height λ, say. Start from a minimum of c, between two L-caps, and suppose the value x_0 of x to be given.

Then x begins to descend. If x_0 is large, x may pass over the top of one or more L-caps, descending all the while, but, sooner or later, it meets an L-cap, passes inside it, and begins to rise. It thereafter meets every cap, for it leaves each cap at a height less than λ, and, since it descends between the caps, it cannot pass over the top of the next. Every year x rises and falls, the maximum occurring later than that of c or L, and we have a yearly epidemic. But there is no reason to make these epidemics equal; though, if any two are equal, they all must be. For if x has equal values for corresponding points of the periods, the whole course following these points is the

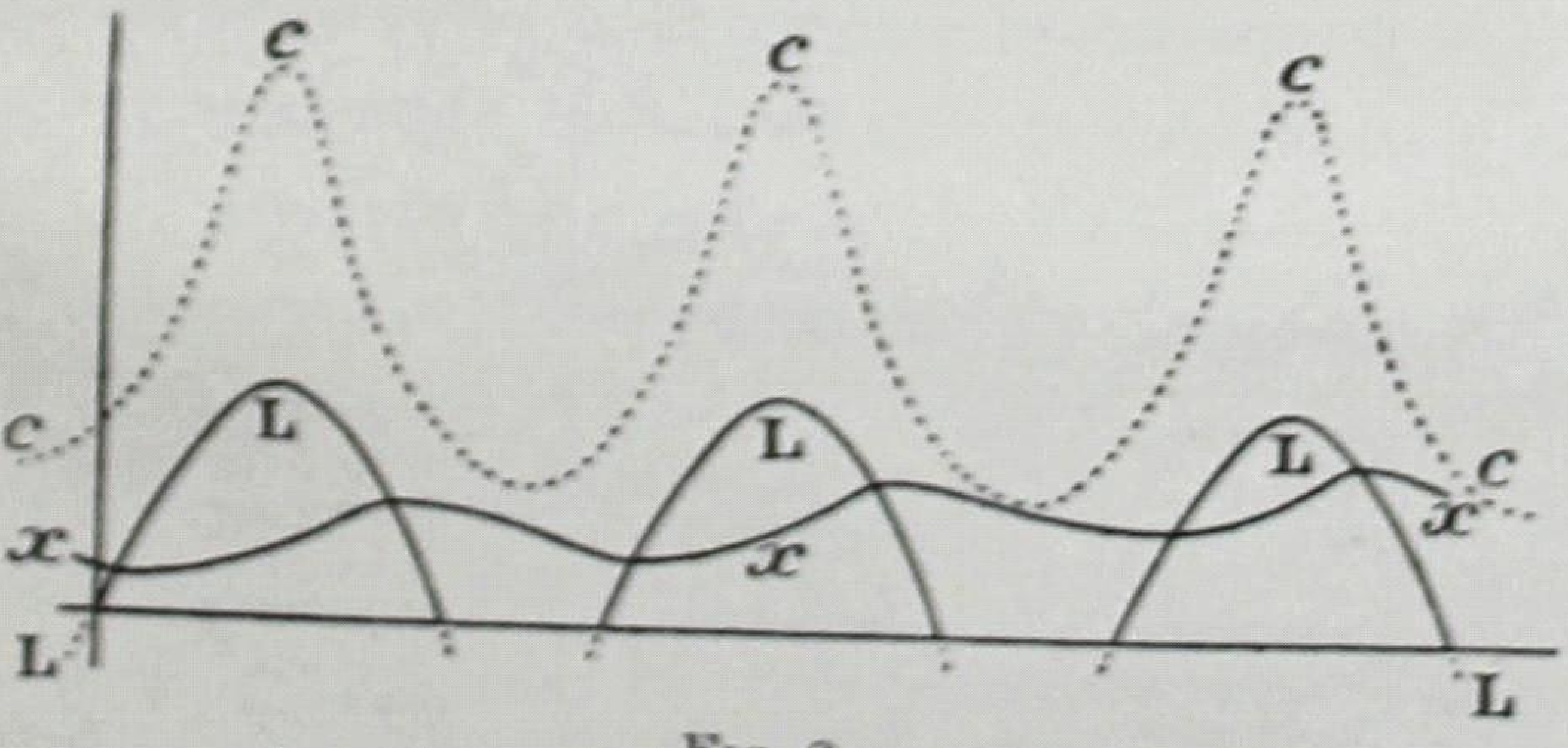

Fig. 2

same; in other words, if two periods are superimposed, and if the two parts of the x-curve have one point in common, they coincide, but, if x meets any L-cap at a point lower than the preceding cap, its whole course inside that cap is lower, and lower still in the next cap, and so on, tending to a limiting form which is periodic, and which may be the axis itself, the happening dying out. But, if x meets any L-cap higher than the last, it meets the next higher still, and so on, and tends to a periodic limiting form of finite fluctuations.

In order to plot x more closely, we have to replace the fundamental differential equation by the approximate difference equation

$$x_{n+1} - x_n = \mathrm{K}x_n (\mathrm{L} - x_n),$$

where K, L are calculated from the value of c measured off the given curve, and x_0 is given; or by some modification of the method.

(v) *Graphical Treatment of f.*—If c is given and x has been obtained, we have next to deal with f, which can be directly calculated from the equation

$$f = cx(1 - x) = c\{\tfrac{1}{4} - (x - \tfrac{1}{2})^2\} = c\{\tfrac{1}{4} - (\tfrac{1}{2} - x)^2\}.$$

Hence f always increases with c: as x increases, f increases if $x < \frac{1}{2}$, and decreases if $x > \frac{1}{2}$. If we wish to know the general behaviour of f, the following Table is useful; it allows for a single bell of each of the c- and x-curves, the maximum of x occurring later than that of c.

c.	x.	f.		
		if $x < \frac{1}{2}$.	if $x = \frac{1}{2}$.	if $x > \frac{1}{2}$.
minimum	decreasing	decreasing	minimum	increasing
increasing	decreasing	?	increasing	increasing
increasing	minimum	increasing	increasing	increasing
increasing	increasing	increasing	increasing	?
maximum	increasing	increasing	maximum	decreasing
decreasing	increasing	?	decreasing	decreasing
decreasing	maximum	decreasing	decreasing	decreasing
decreasing	decreasing	decreasing	decreasing	?
minimum	decreasing	decreasing	minimum	increasing

In using this, we follow f down a column until x passes through the value $\frac{1}{2}$, when we cross to the next f-column.

Since in each column f is increasing in the third row and decreasing in the seventh, there is bound to be a maximum of f between them, which occurs before the maximum of c if $x > \frac{1}{2}$ and after if $x < \frac{1}{2}$.

Recrudescence.—It is seen that if c and x are both decreasing and $x > \frac{1}{2}$, we cannot say what f is doing, as the two causes of its change act in opposite ways. It is possible for f to have a minimum followed by a short rise and a maximum, and, if x then becomes equal to $\frac{1}{2}$, f falls again to another minimum. This would happen, for example, if c suddenly became constant after decreasing, before x drops to $\frac{1}{2}$. Then we have a slight recrudescence of the epidemic near its close.

Since

$$\frac{df}{dt} = \frac{\partial f}{\partial c}\frac{dc}{dt} + \frac{\partial f}{\partial x}\frac{dx}{dt},$$

and in the case considered dc/dt and dx/dt are both negative,

$$\frac{df}{dt} = - x\,(1 - x)\;\left.\frac{dc}{dt}\right| + c\,(2x - 1)\,\mathrm{K}x\,(\mathrm{L} - x)$$

and is positive, so that f rises again, provided dc/dt is small enough, but, as x decreases, the factor $2x - 1$ decreases and changes sign; the positive second term in df/dt becomes negative, and f sinks.

In fig. 3, c rises and falls in a symmetrical bell and then remains constant; L does the same; x has an unsymmetrical bell, rising more steeply than it falls, and sinking to the same limit as L; and f shows an epidemic with slight double recrudescence before sinking to its limit.

The graphs of c, x, and L are on the same scale, and that of f on 10 times the scale.

(vi) f *Assigned.*—The last sub-section started by assuming that c is given in some experimental way. But c is not given except by conjecture, and the method is useful to test whether any particular conjecture leads to numerical results which agree with the records

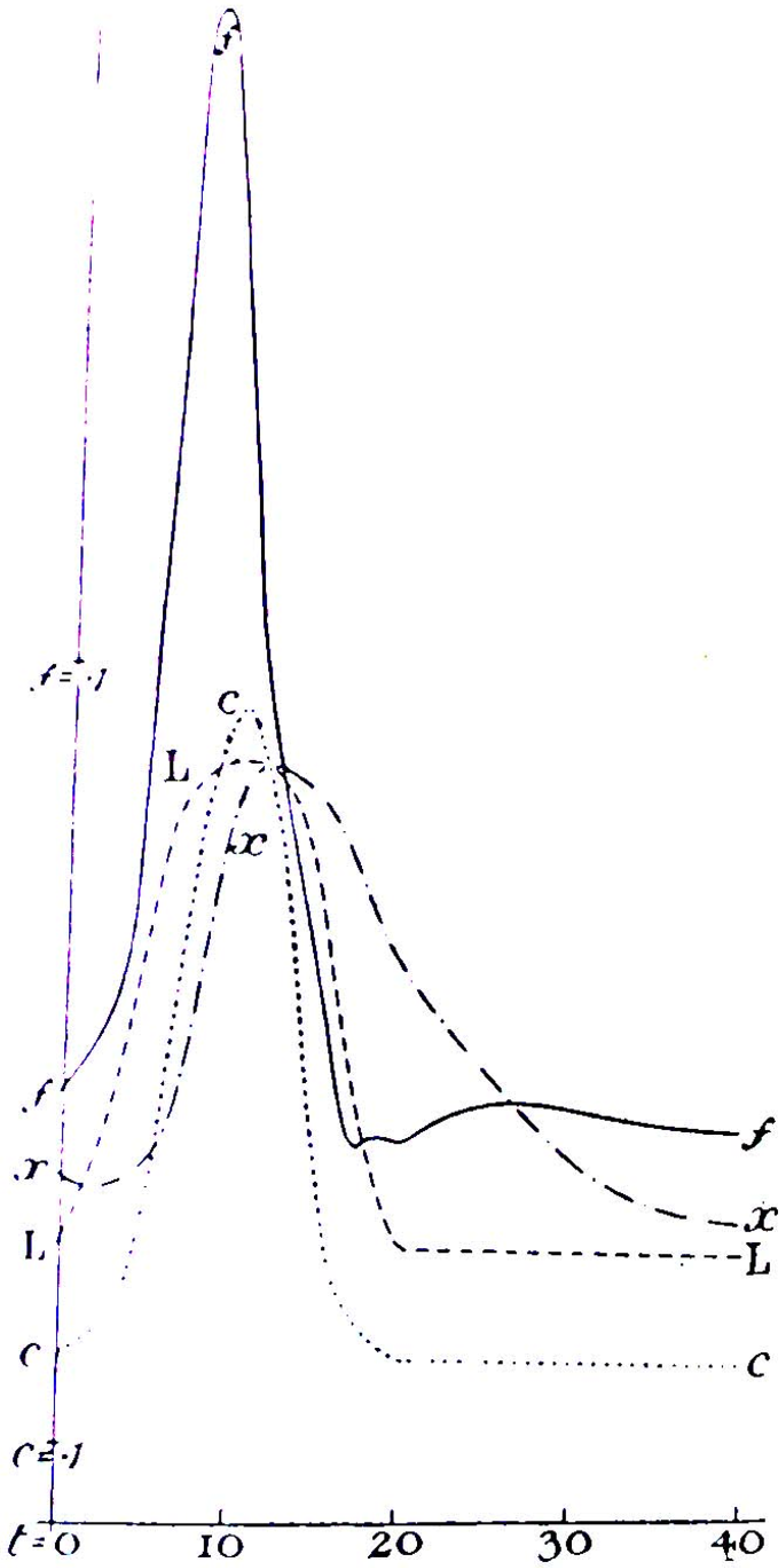
f
f = ·1
c
L
x
f
x
L
c
c = ·1
f
x
L
c
t = 0
10
20
30
40

Fig. 3

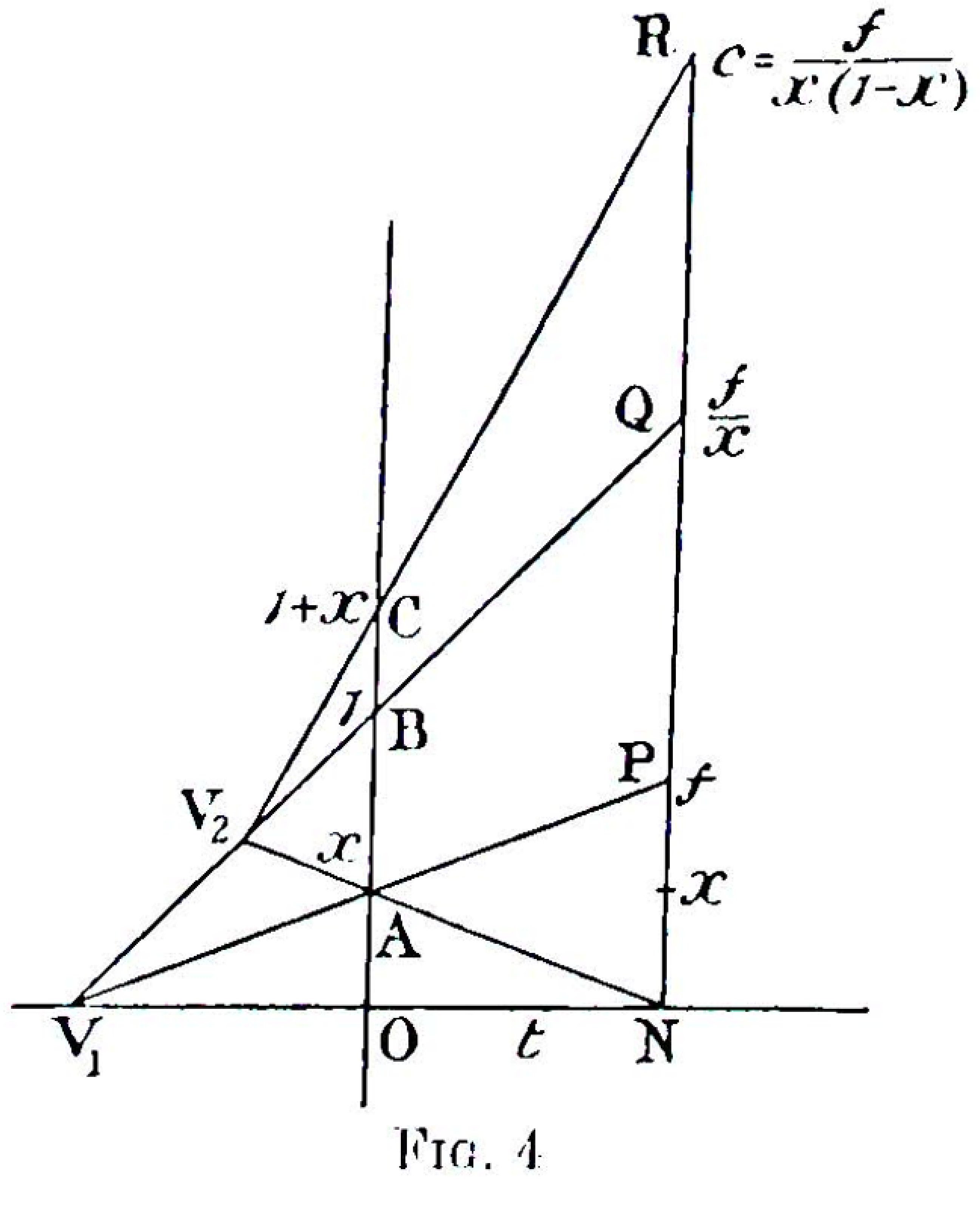
R
$c = \frac{f}{x(1-x)}$
Q
$\frac{f}{x}$
1+x
C
1
B
P
f
V₂
x
x
A
V₁
O
t
N

Fig. 4

of actual epidemics. Now the most extensive and accurate records are the notifications of fresh cases and of deaths made by the medical officers of health: in particular, the records of measles in Aberdeen for the 20 years 1883–1902 are very suitable material, though even there the numbers are not large enough to be quite satisfactory.

The death-rates do not correspond to any of the symbols we have introduced so far; for MZ is the number of deaths of affected persons per unit time, a different thing from the number of deaths due to the affection. We return to this in Section XII (i).

We start therefore with the fresh assumption that the notifications, that is, the values of F, are known. Suppose also that P is given, either by interpolation from census returns, or because in a short and mild epidemic P is nearly constant. Then f is known as a function of t. The fundamental equations are best taken in the forms

$$dx/dt = f - x\,(\mathrm{D} + \mathrm{R} - \mathrm{D}x), \qquad c = f/x\,(1 - x),$$

from which x and c can respectively be determined, if reasonable estimates can be made of the constants D, R.

When x, f are found and plotted, we can construct c graphically. P is the point (t, f). Along the vertical axis mark off

$$\mathrm{OA} = x, \quad \mathrm{OB} = 1, \quad \mathrm{OC} = 1 + x; \quad \text{then} \quad \mathrm{AB} = 1 - x, \quad \mathrm{AC} = 1.$$

Let PA meet the axis NO in V_1: let $\mathrm{V}_1\mathrm{B}$ meet PN in Q and NA in V_2, and let $\mathrm{V}_2\mathrm{C}$ meet PN in R.

Then $\qquad \mathrm{QN}/\mathrm{OB} = \mathrm{PN}/\mathrm{OA}, \qquad \mathrm{QN} = f/x,$

and $\qquad \mathrm{RN}/\mathrm{AC} = \mathrm{QN}/\mathrm{AB}, \qquad \mathrm{RN} = f/x\,(1 - x) = c,$

so that R is the point on the c-curve corresponding to P on the f-curve.

If D and R can both be neglected, then

$$x = \int_0^t \mathrm{F}\,dt + z_0.$$

In this case, x steadily increases whatever is the law of f, and the area under the f-curve always < 1.

In a great many cases, the f-curve is a slightly unsymmetrical bell, rising more steeply than it falls. Now

$$x^2(1-x^2)\frac{dc}{dt} = \frac{df}{dt}x(1-x) - f^2(1-2x),$$

and df/dt is positive for the first, smaller part of the range, then vanishes and is negative for the rest of the time. The second term is negative if $x < \frac{1}{2}$, and positive if $x > \frac{1}{2}$.

Now if x_0 is small, at the summit where $df/dt = 0$, $x < \frac{1}{2}$ and dc/dt is negative. In this case, the c-curve reaches its summit sooner than that of f, and is more unsymmetrical. But if x_0 is larger, $1 - 2x$ changes its sign earlier, and the summit of the c-curve is later in consequence. These results are an extension of those in Part I [p. 28 Section VII (v)], where we found that in the same case where D, R, are neglected, f is perfectly symmetrical when x is constant.

But if we do not wish entirely to exclude reversions, we assume

$$D = 0, \qquad R > 0.$$

Then

$$dx/dt + Rx = f,$$

$$x = \varepsilon^{-Rt}\int_0^t f\varepsilon^{Rt}\,dt + x_0\varepsilon^{-Rt}.$$

For example, let $f = at(T - t)$, which is the simplest function that rises and falls, vanishing at $t = 0$ and $t = T$, with a maximum at $t = \frac{1}{2}T$. Then

$$x = \frac{at}{R^2}\{R(T-t) + 2\} - \frac{a}{R^3}(RT + 2)(1 - \varepsilon^{-Rt}) + x_0\varepsilon^{-Rt}.$$

Or we may allow for asymmetry by an exponential factor; $f = at(T - t)\,\varepsilon^{-\lambda t}$, which still vanishes at $t = 0$ and $t = T$, but has a maximum before $t = \frac{1}{2}T$;

$$x = \frac{at}{(R-\lambda)^2}\{(R-\lambda)(T-t) + 2\}\varepsilon^{-\lambda t}$$
$$- \frac{a}{(R-\lambda)^3}\{(R+\lambda)T + 2\}(\varepsilon^{-\lambda t} - \varepsilon^{-Rt}) + x_0\varepsilon^{-Rt}.$$

If f is periodic, let

$$f = a + b \sin pt, \qquad \text{where } a \geqq b\,;$$

$$x = \frac{a}{\mathrm{R}}(1 - \varepsilon^{-\mathrm{R}t}) + \frac{b}{p^2 + \mathrm{R}^2}(\mathrm{R} \sin pt - p \cos pt + p\varepsilon^{-\mathrm{R}t}) + x_0\varepsilon^{-\mathrm{R}t}.$$

In all these, the negative exponentials decrease rapidly ; in the last, x oscillates about a/R as mean, the phase lagging behind that of f by $\frac{1}{p} \tan^{-1} \frac{p}{\mathrm{R}}$. Then c has the same period.

(vii) *More General Equations.*—In many diseases, for example, those in which there is a period of incubation, the infectivity is variable, and depends not only on the absolute time but also on the *duration of the case*, which is the time elapsed since the case was first infected. In order to take this into account, c must be considered as a function of two independent variables, t the time, and s the duration ; fluctuations of the infectivity of the disease as a whole cause c to vary as a function of t ; episodes in the course of a single case cause c to vary as a function of s, very possible a discontinuous function. For example, an incubation period of q days is represented by taking $c = 0$ when $s < q$.

For greater generality, suppose that V, I, E, M, N, c, r, D, R, all depend on t and s ; but v, i, e, m, n on t only.

Let $\mathrm{F}_{t,\,0}\,dt$ be the number of fresh cases occurring between times t and $t + dt$.

Let $\mathrm{F}_{t+s,\,s}\,dt$ be the number of the batch of $\mathrm{F}_{t,\,0}\,dt$ cases which survive, having neither died, emigrated, nor reverted, at time $t + s$, being then of duration s. Thus $\mathrm{F}_{t,\,s}$ is the number of cases existing at time t whose duration is s. The total number of cases existing at time t is

$$\mathrm{Z} = \int_{s=0}^{\infty} \mathrm{F}_{t,\,s}\,ds. \tag{86}$$

As before, $\mathrm{P} = \mathrm{Z} + \mathrm{A}$, and the number of fresh cases is now

$$\mathrm{F}_{t,\,0} = \frac{\mathrm{A}}{\mathrm{P}} \int_{s=0}^{\infty} c\mathrm{F}_{t,\,s}\,ds.$$

The rate of change of the particular batch of cases which started at time t_0 is given by

$$\frac{d}{dt} \mathrm{F}_{t,\, t-t_0} = (\mathrm{V} - \mathrm{N} - r)\, \mathrm{F}_{t,\, t-t_0},$$

which may be written

$$\left(\frac{\partial}{\partial t} + \frac{\partial}{\partial s}\right) \mathrm{F}_{t,} = (\mathrm{V} - \mathrm{R})\, \mathrm{F}_{,\, s}, \tag{87}$$

where

$$s = t - t_0.$$

The fundamental equations now take the forms

$$\left.\begin{aligned} \frac{d\mathrm{P}}{dt} &= v\mathrm{A} + \int_{s=0}^{\infty} \mathrm{VF}_{t,\, s}\, ds \\ \frac{d\mathrm{Z}}{dt} &= \int_{s=0}^{\infty} (\mathrm{V} - \mathrm{R})\, \mathrm{F}_{t,\, s}\, ds + \mathrm{F}_{t,\, 0} \\ \frac{d\mathrm{A}}{dt} &= v\mathrm{A} + \int_{s=0}^{\infty} \mathrm{RF}_{t,\, s}\, ds - \mathrm{F}_{t,\, 0} \end{aligned}\right\}. \tag{88}$$

These equations are not independent. If 87 is integrated with respect to s from 0 to ∞, since $\mathrm{F}_{t,\, \infty} = 0$ we have

$$\frac{d}{dt} \int_{s=0}^{\infty} \mathrm{F}_{t,\, s}\, ds - \mathrm{F}_{t,\, 0} = \int_{s=0}^{\infty} (\mathrm{V} - \mathrm{R})\, \mathrm{F}_{t,\, s}\, ds,$$

which is equivalent to 88 by virtue of 86.

These equations as they stand are quite unwieldy, but they are capable of being applied to a large number of simpler cases.

If V, R, are independent of t and constant for a particular range of s, say $s = q_1$ to q_2, then for any intermediate value of s, from 87,

$$\mathrm{F}_{t,\, s} = \mathrm{F}_{t-q_1,\, q_1}\, \varepsilon^{(\mathrm{V}-\mathrm{R})(t-q_1)},$$

and if V, R have other constant values V′, R′ from $s = q_2$ to q_3, then for any value of s in the latter range,

$$\begin{aligned} \mathrm{F}_{t,\, s} &= \mathrm{F}_{t-q_2,\, q_2}\, \varepsilon^{(\mathrm{V}'-\mathrm{R}')(t-q_2)} \\ &= \mathrm{F}_{t-q_1,\, q_1}\, \varepsilon^{(\mathrm{V}'-\mathrm{R}')t+(\mathrm{V}-\mathrm{V}'-\mathrm{R}+\mathrm{R}')q_2-(\mathrm{V}-\mathrm{R})q_1}. \end{aligned}$$

and so on.

The simplest case is when c is constant from $s = q_1$ to q_2, zero for all other values of s, and independent of t, while all the other elements are neglected. Since there are then no deaths nor recoveries, $F_{t,s} = F_t$ and P is constant.

Then

$$\frac{dZ_t}{dt} = F_t = e\frac{A}{P}\int F_{t-s}\,ds$$

$$= e(1 - Z_t/P)(Z_{t-\iota} - Z_{t-\lambda}),$$

the suffixes showing the value to be substituted for t in Z.

Let $e(1 - Z_t/P) = Y_t$, then this difference equation takes the simpler form

$$\frac{dY_t}{dt} = Y_t(Y_{t-\iota} - Y_{t-\lambda}).$$

The best way of dealing with this equation is to assume a Dirichlet series for Y, substitute and equate coefficients.

Let $Y_t = B_0 + B_1\varepsilon^{-pt} + B_2(B_1\varepsilon^{-pt})^2 + B_3(B_1\varepsilon^{-pt})^3 + \ldots$
and let $\Delta_n = \varepsilon^{n\iota p} - \varepsilon^{n\lambda p}$

then p, B_1 are arbitrary, $B_0 = p/\Delta_1$, and the general reduction formula for the coefficients is

$$B_n p(\Delta_n - n\Delta_1)/\Delta_1$$
$$= B_{n-1}\Delta_1 + B_{n-2}B_2\Delta_2 + B_{n-3}B_3\Delta_3 + \ldots + B_{n-1}\Delta_{n-1}.$$

Since $\Delta_n - n\Delta_1$ is positive, the signs of the coefficients alternate $B_0 = Y_\infty$ and is essentially positive, and therefore p is positive;

$$B_1 = -\frac{1}{p}\frac{dY_t}{dt} = \frac{e}{Pp}\frac{dZ_t}{dt}$$

and is positive, for Z always increases. Therefore B_{2n} is negative and B_{2n+1} positive.

The series for F_t is

$$F_{t,0} = \frac{pP}{e}\{B_1\varepsilon^{-pt} + 2B_2(B_1\varepsilon^{-pt})^2 + 3B_3(B_1\varepsilon^{-pt})^3 + \ldots\}.$$

SECTION XII

(1) *Mortality: First Modification of the Equations of Part I.*—It has been assumed that the average death-rate of the affected popula-

tion is a constant M; part of this, m say, is due to other causes, and $M - m$ is due to the disease. If m is constant, so is $M - m$, and the deaths from disease are simply proportional to Z and their ratio to the total population P is proportional to x. It therefore increases steadily; this does not agree with the records.

It is more reasonable to assume that the deaths from the disease all occur soon after infection and that their number is proportional to F; let it be μF. Then its ratio to the total population is μf, and rises and falls in proportion to f. This is much more nearly true. But now the fundamental equations are altered. We assume that the μF deaths are in addition to MZ deaths due to other causes, and that they all happen before the cases have had time to infect an appreciable number of other people. Then

$$dP/dt = VZ + vA - \mu F,$$
$$dZ/dt = (V - N - r)\,Z + (1 - \mu)\,F,$$
$$F = cZA/P.$$

As before write

$$Z = xP, \quad A = (1 - x)\,P, \quad F = cx\,(1 - x)\,P,$$
$$v - V = D, \quad N + r = R.$$

The result of elimination is

$$\begin{aligned} dx/dt &= -\,\{D + R - (1 - \mu)\,c\}\,x \\ &\qquad + \{D - (1 - 2\mu)\,c\}\,x^2 - \mu c x^3 \\ &= -\,\mu c x\,(x - \alpha)\,(x - \beta), \text{ say, where } \alpha > \beta. \end{aligned}$$

of which the integral is

$$\varepsilon^{-\mu c \alpha\beta\,(\alpha - \beta) t} = x^{\alpha - \beta}\,(x - \alpha)^{\beta}\,(x - \beta)^{-\alpha} / \{x_0^{\alpha - \beta}\,(x_0 - \alpha)^{\beta}\,(x_0 - \beta)^{-\alpha}\}.$$

The form of this curve depends on the signs of α, β.

Let
$$\begin{aligned} Q &= \mu c\,(x - \alpha)\,(x - \beta) \\ &= \mu c x^2 - \{D - (1 - 2\mu)\,c\}\,x + D + R - (1 - \mu)\,c. \end{aligned}$$

Case (i).—If the last term is positive, $(1 - \mu)\,c < D + R$, then α, β, have the same sign; and if x_0 is a small positive quantity, less than $|\alpha|$ or $|\beta|$, dx_0/dt is negative, x begins to decrease, Q cannot change sign and x decreases steadily to zero as its limit.

Case (ii).—If the last term in Q is negative, $(1 - \mu) c > D + R$, then α is positive and β is negative. When $x = 1$, $Q = R$ and is positive; therefore $0 < \alpha < 1$. Now dx_0/dt is positive, x begins to increase, but it does not reach α till $t = \infty$; hence Q cannot change sign and x increases steadily to α as its limit.

Now $f = cx(1 - x)$, and as x varies from 0 to 1, f varies from 0 to 0 in a bell-shaped curve with a maximum at $x = \frac{1}{2}$.

In case (i) f decreases steadily from f_0 to 0, in a curve of the same general character as the x-curve, for the factor $(1 - x)$ does not vary much.

In case (ii) if $\alpha < \frac{1}{2}$, f increases steadily to $c\alpha(1 - \alpha)$ as its limit but if $\alpha > \frac{1}{2}$, f has a maximum $\frac{1}{4}c$ when $x = \frac{1}{2}$ and then sinks to its limit $c\alpha(1 - \alpha)$.

If D, R are both negligible, $Q = c(x - 1)(\mu x + 1 - \mu)$, and $\alpha = 1$; then f has a maximum and sinks to 0 as its limit.

(ii) *Second Modification.*—A popular assumption about infectious diseases is that the case ceases to be infective after a certain time, when also there ceases to be danger of death from the disease, but remains immune permanently. To represent this, suppose the population P to consist of three classes: $P = A + X + Y$, where A are unaffected, X infective and immune, Y immune. For simplicity, neglect all "natural" variations, but suppose there are MX deaths from the disease per unit time, rX recoveries, and $F = cXA/P$ fresh cases. The fundamental equations are

$$dP/dt = -MX, \qquad dX/dt = cXA/P - (M + r)X,$$

$$dY/dt = rX, \qquad dA/dt = -cXA/P.$$

Hence P decreases and Y increases steadily. Suppose $Y_0 = 0$. Then

$$dY/dP = -r/M, \qquad Y = (r/M)(P_0 - P);$$

$$dA/dP = c/M \,.\, A/P, \qquad A = A_0(P/P_0)^{c/M};$$

$$-(1/M)dP/dt = X = (1 + r/M)P - (r/M)P_0 - A_0(P/P_0)^{c/M};$$

$$t = -\frac{1}{M}\int_{P=P_0}^{P} dP/\{(1 + r/M)P - (r/M)P_0 - A_0(P/P_0)^{c/M}\};$$

which is integrable in finite form if c/M is rational, and leads to algebraic functions and logarithms.

In the final steady state, $\mathrm{X} = 0$ and P is given by

$$(1 + r/\mathrm{M})\,\mathrm{P} - (r/\mathrm{M})\mathrm{P}_0 - \mathrm{A}_0(\mathrm{P}/\mathrm{P}_0)^{c/\mathrm{M}} = 0,$$

which has a real root P_∞ between P_0 and 0. A definite number of persons $\mathrm{A}_\infty = \mathrm{A}_0\,(\mathrm{P}_\infty/\mathrm{P}_0)^{c/\mathrm{M}}$ escape the disease altogether.

Now, $d\mathrm{X}/dt$ is of the same sign as $c\mathrm{A} - (\mathrm{M} + r)\mathrm{P}$, and is positive at first, provided $c > (\mathrm{M} + r)\mathrm{P}_0/\mathrm{A}_0$; in this case X, the total number of infective persons, rises to a maximum before sinking to 0.

dx/dt is found to be of the same sign as $(c - \mathrm{M})\mathrm{A} - r\mathrm{P}_0$, which is positive at first if $c > \mathrm{M} + r\mathrm{P}_0/\mathrm{A}_0$, and steadily decreases.

$d\mathrm{F}/dt$ is of the same sign as $c\mathrm{A} - (\mathrm{M} + r)\mathrm{P} - (c - \mathrm{M})\mathrm{X}$, and, if $c > \mathrm{M}$, this is less than $d\mathrm{X}/dt$; the maximum of F (if it has one) is therefore earlier than that of X or of MX, that is to say, the greatest number of deaths occurs later than the height of the epidemic, which is the time of greatest number of notifications.

For example, let

$$r = 0{\cdot}01, \qquad \mathrm{M} = 0{\cdot}02, \qquad c = 0{\cdot}04\,, \qquad x_0 = 0{\cdot}0102.$$

Then $\mathrm{A}_0/\mathrm{P}_0 = 0{\cdot}9898.$ Let $\mathrm{P}/\mathrm{P}_0 = p$,

$$\mathrm{M}t = \int_{p=1}^{p} \frac{dp}{0{\cdot}9898p^2 - 1{\cdot}5p + 0{\cdot}5}$$

$$= \frac{-1}{0{\cdot}52} \log_\epsilon \left\{ \frac{2{\cdot}02p - 1}{0{\cdot}98p - 1} \cdot \frac{-0{\cdot}02}{1{\cdot}02} \right\}$$

$$p = \frac{1 + 51\varepsilon^{-0{\cdot}0104t}}{2{\cdot}02 + 49{\cdot}98\varepsilon^{-0{\cdot}0104t}}$$

$\mathrm{P}_\infty = \frac{1}{2}\mathrm{P}, \qquad \mathrm{A}_\infty = \frac{1}{2}\mathrm{P}_\infty$ approximately.

X is stationary when $\mathrm{A} = \frac{3}{4}\mathrm{P}$ or $1/p = 1{\cdot}32$, leading to $t = 300$ approximately, the time which elapses before the maximum of the mortality.

SECTION XIII

(i) *Discussion.*—In the foregoing work, very little attention has been paid to the actual values of the constants. We have rather been preparing apparatus which can be used in a great variety of special cases.

It will be seen that the number of dependent variables is large, and that a fluctuation in any one can be accounted for in a great many different ways by supposing that there is a suitably adjusted variation in almost any one of the parameters. And the variation of any one parameter when it is over a range small compared with the mean value can be represented approximately by one of several analytic functions, which would lead to the most diverse results if the variation were to continue over a wider range. It is therefore extremely rash to rule out *a priori* any one possible cause of fluctuation of any one phenomenon.

The cases already considered have led exactly to the series of curves required by the facts :

1. The steadily rising curve of a happening that gradually permeates the whole population [VII (iii)].
2. The symmetrical bell-chaped curve of an epidemic that dies away entirely [VII (v)].
3. The unsymmetrical bell of a new happening that begins with an epidemic, and settles down to a steady endemic level [VII (v)].
4. The periodic curve with regular rise and fall due to seasonal disturbances [XI (iv)].
5. The more irregular curve where there is recrudescence before the end of an epidemic, or where outbreaks differing in violence occur at unequal intervals [XI (v)].

This suggests that, as was stated in the Introduction to Part I, the rise and fall of epidemics as far as we see at present can be explained by the general laws of happenings, as studied in this paper.

The methods which have been used call, however, for the following remarks :—

A. In the equations of Section VII, the case-mortality and infectivity were both regarded as constants. It is very unlikely that this is strictly true ; and it is much more probable that their values are highest at some definite short period after infection and then gradually die away.

B. In Section XI some attempt is made to meet this objection by introducing the case-duration as a second independent variable, in addition to the time, but there are innumerable other considerations that may also be taken into account, such as the age, sex, and social position of the patients, climate, temperature, and, in the case of insect-borne diseases, the habits of the carrier.

C. Except in XII (ii), the whole population has been divided into two classes only, the unaffected, and the affected who are considered to be both infective and also immune. But it is again unlikely that this is strictly true, and, in order to represent the facts accurately, a much larger number of classes would have to be considered.

D. Although it would probably be easy to write down equations on any given hypothesis as to the number and behaviour of the variables, these equations, like those of XI (vii), would probably be quite unsuited to numerical work, and we should have to fall back on laborious approximate methods such as expansion in series or mechanical quadratures.

E. The quantities x and L have been treated as continuous functions of the time, but, as they represent numbers of persons, they can only take integral values, and must be discontinuous unless they are constant; in the same way, f and l can only vary by integral multiples of 1/P. This is a serious drawback when the numbers are not large. In a small village, the average number of cases in the period between two epidemics may be only two or three, and, if these are removed by any chance, the whole course of events is entirely changed; there is no one to keep alive the infection, and a new generation may grow up entirely non-immune. Then the entry of a fresh source of infection has the same effect as if c rose suddenly from 0 to its normal value. In fact, in the case of small populations, other than the most probable values may often obtain, so that the whole of this book is applicable to large numbers rather than to small ones.

Harrison and Sons, Ltd., Printers, St. Martin's Lane, London, W.C.2.

17337
16-x-57

ALLAMA IQBAL LIBRARY
17337